EFFECTIVENESS OF A MICROCOMPUTER-BASED LABORATORY IN LEARNING DISTANCE AND VELOCITY GRAPHS

By

HEATHER MARIANNE BRASELL

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN
PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

1987

ACKNOWLEDGEMENTS

I am indebted to many people and organizations for the use of facilities and equipment for this research. Technical Education Research Centers (TERC), Cambridge, MA, were generous in providing use of MBL equipment, software, worksheets, and research data, as well as unlimited access to developers and researchers for advice and assistance. They also provided the means for me to participate in and to present my research findings and ideas to a symposium in Cambridge. I am very grateful to Dr. Robert Tinker, Dr. Janice Mokros, Dr. Ronald Thornton, and Dr. Tim Barclay from TERC for their support, encouragement, and assistance throughout my research.

Dr. Sharon Dugdale and Dr. Dave Kibbey, of University of Illinois, allowed me to use their software in the research, and very generously adapted it for my requirements, transforming it from an instructional program to a test. Dr. Mary Budd Rowe lent me her personal computer, very much easing the logistics of working in small rural schools. University of Florida provided many facilities that were invaluable for my research: Center for Instructional and Research Computing Activities (computing facilities and consultation), Educational Media Center (for loan of audiotaping and videotaping equipment), and the many university libraries with the interlibrary loan service.

I spent many hours with the two physics classes at P. K. Yonge Laboratory School as they worked through the mechanics section of their

course. I am very grateful to the school and students, and particularly to Dr. Paul Becht, for allowing me to observe and participate so freely in the class activities, and to pilot test the MBL activities. I learned a great deal from both students and teacher. I would also like to thank the principals and county school superintendents who allowed me to conduct my research in the nine schools, and the students and teachers who participated in the study for their willingness and cooperation.

I am grateful to Dr. Steve Pinker (Massachusetts Institute of Technology), Dr. Barbara White (Bolt, Beranek, and Newman), Dr. Michael Padilla (University of Georgia), and Dr. Marcia Linn (University of California, Berkeley) for sending me unpublished and prepublished information; to Steve Beardsley and Dr. Elroy Bolduc for advice in adapting software; Dr. Robert Soar for advice in constructing my observation instrument and for statistical advice; to Dr. Murray Gaskins and TERC personnel for constructing hardware; to my supervisory committee of Dr. Mary Budd Rowe (chair), Dr. Elroy Bolduc, Dr. Eugene Chenette, and Dr. Robert Soar for comments and guidance throughout my research and dissertation preparation; and to Dr. Parker Small and Dr. Marcia Linn for extensive comments on manuscripts in preparation. Throughout my graduate studies, Dr. Mary Budd Rowe has stimulated my ideas and my ideals in the area of science education, and has provided some invaluable opportunities to test and apply these ideas.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Abstract of Dissertation Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

EFFECTIVENESS OF A MICROCOMPUTER-BASED LABORATORY
IN LEARNING DISTANCE AND VELOCITY GRAPHS

By

HEATHER MARIANNE BRASELL

May, 1987

Chairman: Mary Budd Rowe
Major Department: Instruction and Curriculum

Microcomputer-based laboratories (MBL) use the computer to collect data, to transform the data, and to display it on the screen. This research examines the impact of a single class period of experience with the MBL motion unit on high-school students' ability to develop a cognitive linkage between a physical movement event and the Cartesian graph of either distance or velocity displayed on the computer screen. By focusing specifically on finding out know how and why MBL activities are effective in improving students' graphing skills (i.e., attributes of both MBL as tool and students as learners), this research complements studies of extended MBL experience.

The impact of real-time graphing, selected as the most salient feature of the MBL tool, was isolated by delaying the graph display for 20-30 seconds but leaving the activity otherwise identical. These two

experimental treatments were compared with a pencil-and-paper graphing activity and a pretest/posttest treatment. A factorial analysis of covariance determined that students in the standard-MBL treatment group had significantly ($p<.05$) lower error rates on the posttest than did students in either the delayed-MBL or the pencil-and-paper treatments. When the posttest was subdivided, the same pattern was found with distance-time items, but not with velocity-time items. Improvement seems to have been restricted to understanding of the concepts of both distance (not velocity) and graphs. Most of the posttest differences could be attributed to the immediacy of graph display in the standard-MBL treatment. Other features of the real-time MBL, such as "dynamic" display, increased opportunity for practice, and changes in management of laboratory time, did not generate any improvement in the graphing skills measured in this study when compared with the pencil-and-paper treatment.

Observations of students' behaviors indicated substantial differences in the quality of cognitive engagement and motivation of the two MBL treatment groups. The MBL activity provided students with an opportunity for real experimentation and data that were worth interpreting. The responsiveness of the real-time graphing seemed to improve motivation and provided a sense of competence and achievement. Some sex differences in attitude and performance were revealed.

Students' performance on a number of measures indicated a superficial competence with graphing. Almost one fifth of students had so much difficulty in constructing a graph from raw data that they appeared to lack a functional graph schema.

CHAPTER I

INTRODUCTION

Microcomputer-Based Laboratories

Microcomputer-based laboratories (MBL) are laboratories where students use computers, as laboratory tools, to acquire and analyze experimental data. They can measure a variety of phenomena, such as sound, light, motion, temperature, and time. They can detect and measure these properties over time scales that were previously impossible or impractical (either shorter or longer). Computers can be programmed to obtain the information quickly and accurately, and users can see the information displayed on the computer screen as it is being measured, without having to perform the time-consuming, tedious procedures of calculations and manual graph construction. That is part of the computer program. It gives students an unprecedented ability to to explore their environment.

However, not all MBLs are equivalent. Various approaches have been followed in developing MBL protocols. Some developers voice concerns about the "black box" syndrome. "Any use of sophisticated equipment in a teaching laboratory . . . carries with it the danger of the 'Black Box Syndrome;' that is the student may lose sight of the laws under

investigation and learn only how to manipulate the dials of the black box" (McInerney, 1984, p. 20). Such developers require students to write the software and to graph the data. Some teachers have similar reservations, believing that students need to (a) manually construct graphs in order to develop graphing skills, and (b) manually perform the calculations in order to reinforce the mathematical relationships.

The equipment and software used in this dissertation were developed by Technical Education Research Centers (TERC) in Cambridge, Massachusetts, who hold that having the computer perform the time-consuming tasks of calculating and constructing graphs enables students to concentrate on observing the phenomena under investigation, rather than on performing the mechanical acquisition and manipulation of data. Students control the computer, and the computer performs the routine procedures of collecting measurements, performing calculations, and displaying these data via a graph on the computer screen. In short, data collection and display are automated. Students control several key aspects of the activity: (a) they decide what to measure, (b) they select scales and parameters for collecting and displaying data, and (c) they interpret the results.

Two key features distinguish MBLs from traditional laboratories. The first is real-time graphing--the data being measured, and/or any data being derived from these measurements, are displayed in real time (i.e., at the same time that they are acquired, one datum point at a time). The students can see the graph appearing dynamically on the computer screen, reflecting changes in the physical system.

The second feature is time management. Because MBL activities relieve students of the physical labor of taking measurements, performing calculations, and constructing tables and graphs to report their data, the focus of laboratory activities shifts from collecting data to examining and interpreting data. Observations in MBL classrooms indicate that time saved is spent on three kinds of activities related to the objectives of laboratory experiences: (a) doing more activities that converge on the central concept, (b) replicating both to improve and to verify results (both important to training), and (c) generating and testing hypotheses. Such activities should provide students with the opportunity to develop a sound understanding of the basic processes of science that ordinarily do not get much exercise in the noncomputer-based condition (Hegarty, 1982).

Research Focus

Status of MBL Research

There are a number of practical and educational reasons why this new development in science education is seen to be attractive--practical because it saves time, and educational because it induces some new behaviors on the part of students. Findings from preliminary field trials and research (Barclay, 1986; Linn, Layman, & Nachmias, in press; Mokros, 1986; Mokros & Tinker, in press; Thornton, 1986) indicate that use of MBL protocols improves cognitive performance and motivation for

students from sixth grade to college. Students appear to improve their understanding in three general areas: (a) science concepts, (b) graphs as a symbolic representation of these concepts, and (c) basic science processes. As these kinds of laboratories focus student attention on the meaning and use of graphs to interpret the properties of physical phenomena, they may be particularly powerful in assisting students to understand graphs as representations of events (Mokros, 1986; Mokros & Tinker, in press; Thornton, 1986). If so, this would have considerable educational significance because graphs are so important as a medium of thinking, learning, and communicating in science.

Preliminary research findings have partially substantiated these claims, but this has been mainly for middle-school students (Barclay, 1986; Mokros, 1986; Thornton, 1986). More recently, two studies (Linn et al., in press; Mokros & Tinker, in press) have demonstrated improvements with performance in interpreting graphs following extended experience (several weeks) with MBL activities. Linn et al. (in press) have compared these improvements with some comparison data from non-MBL control groups. However, in both studies, about one third of the MBL students continued to make fundamental errors in interpreting graphs, even when the content of the graphs was the same as for their MBL experience. Research needs to address this issue of persistence in errors and to ask more specific questions about what skills are facilitated by experience with MBL.

It is important that we ask why and how MBLs achieve these improvements. Based on observations from field trials, developers suggest that the real-time graphing feature is important to the learning

that occurs. They claim that immediate graphic reinforcement of abstract concepts helps students to develop a feel for the physical property being studied. "Preliminary observations indicate that the linking of concrete measurement of an actual physical system with the simultaneous production of the symbolic representation may be an effective way for students to learn to correctly interpret and produce graphs" (Thornton, 1986, p. 1). One way of determining which specific features of MBL are responsible for any facilitative effects is to identify key features of MBL that differ from more conventional laboratory techniques, to experimentally separate these features under different treatment conditions, and thus to examine how each facilitates specific types of information processing.

Dissertation Research Agenda

The research was designed to complement the research described above, by (a) studying a different student population, high school physics students, (b) using a very brief treatment period to focus on the initial impact of MBL with minimal instructional support, (c) examining in finer grain specific attributes of the MBL tool, and (d) exploring attributes of the learner that are likely to help or hinder students in developing the cognitive linkage between a physical event and a graph.

The research agenda of this dissertation addressed four fundamental questions.

1. Are microcomputer-based laboratories (MBLs) more effective than

conventional techniques that rely on manual data collection, calculation, and graphing (specifically a pencil-and-paper, graph-construction activity)

a) for learning to interpret graphs, and

b) for learning concepts?

These skills were measured by the ability to translate between verbal and graphic descriptions of physical events.

2. If the MBL activity is effective for learning these things, what specific feature(s) of the MBL makes it effective for learning both the physics concepts and graphing skills? In this study, it was possible to consider only one such attribute--real-time graphing of data. Other important features of MBL, not considered separately, include

   a) dynamic graphing of data,

   b) graphing performed by computer rather than students, and

   c) increased opportunity for practice or repetition.

3. What kinds of problems do students have with graphs--in understanding, interpreting, and constructing graphs?

4. What kinds of students have problems with graphs? Gender was one attribute of particular interest.

The research described in this dissertation was focused predominantly on the first two of these questions as they relate to a specific MBL activity using the motion detector. It examined whether, how, and why the MBL motion detector helps students to learn about distance-time and velocity-time graphs. These concepts constitute a perennial problem for physics teachers and students.

To address the relevant research questions, the research goals were threefold. First, at a pragmatic, pedagogic level, using a very brief treatment period provides empirical evidence for the "first-strike" impact of MBL in facilitating students in learning to translate between (link) verbal and graphic descriptions of a physical event. Second, isolating a single salient feature of MBLs--real-time graphing--and examining its contribution to the improved students' performance starts the process of determining how and why MBLs succeed in improving graphing skills. This was done by contrasting two treatments, one of which had the graphed data displayed in real time, whereas the other delayed the display of the data for a short time (20-30 seconds). Finally, examining performance on a range of tests of graphing skills should contribute to our understanding of students' graphing skills as both a dependent and an independent variable.

Treatments

Research in this dissertation concentrated on determining the impact of real-time graphing as the most salient feature of the MBL tool. The effect of real-time graphing of the standard MBL was isolated by delaying the graph display until all the data had been collected, a delay of 20-30 seconds, but leaving the activity otherwise unchanged. These two experimental treatments, standard MBL and delayed MBL, were compared with a pencil-and-paper, graph-construction activity as a reference activity treatment, and with a no-activity treatment as a statistical control.

Research Hypotheses

The experimental hypotheses derive from the claims of developers and the experience of teachers who suggest that much of the effectiveness of MBL resides in the real-time graphing feature. If that were so, then for high school physics students, real-time graphing of data from an experimental event during a kinematics MBL activity, would significantly improve the development of a cognitive connection between the physical phenomenon and its representation on a line graph compared with activities where the same graph display is delayed, and in comparison with a pencil-and-paper activity. The development of this cognitive connection is assessed by reduction in error rate in translating between verbal and graphic descriptions of simple motion events.

However, that hypothesis is only one of three alternative models or sets of expectations. The second alternative hypothesis recognizes that any improvement in students' performance may be due to features other than real-time graphing. If that were the case, then we would expect that performance on the category of tasks described above would be better after delayed-graph experience than it would be after the pencil-and-paper treatment activity.

The third alternative hypothesis acknowledges the theoretical and research bases that suggest that information processing during a short delay may enhance learning and remembering. If that were so, then performance after delayed-graph treatment could be better than after real-time graphing. However, whether the real-time or delayed graphing

proves more effective may depend on the type of information processing undertaken and the motivation of the students. This discussion will be elaborated in the literature review.

Statistical Hypotheses

The research hypotheses were translated into three specific null hypotheses for testing. Given the nature of the alternative hypotheses, directional statistical hypotheses were not appropriate. In brief, the statistical hypotheses were that there would be no significant differences ($p<.05$) on posttest measures between (a) the standard-MBL treatment and the delayed-MBL treatment, (b) the standard-MBL treatment and the paper-and-pencil treatment, and (c) the delayed-MBL treatment and the pencil-and-paper treatment. These hypotheses will be elaborated further and explained in Chapter III.

Definitions, Limitations, and Assumptions

The terms of the hypotheses need to be qualified and defined. These restrictions, which limit the generalizability of the results, will be elaborated in the methods chapter.

1. Real-time graphing of data indicates a procedure in which data are generated from a physical event and displayed immediately in graph format, as the event is occurring. Collecting data, calculating, and graphing procedures are automated by the computer and sensor.
2. Delayed graphing of data indicates that data are displayed in graph

format after a delay of about 20-30 seconds (i.e., the event and the graph occur sequentially).

3. The MBL hardware and software were developed by Technical Educational Research Centers (TERC). The specific probe was the sonic rangefinder, or motion detector. The software permits the data to be displayed as either a distance- or velocity-time graph.
4. The MBL activity was a single-class-period, inductive, small-group activity where both teacher and experimenter were minimally involved. Guidance was provided by worksheets.
5. The content of the laboratory unit was the kinematics concepts of distance and velocity, but not acceleration.
6. The graphs used as the medium of communicating data were line graphs within a Cartesian coordinate system.
7. The student population consisted of high school physics students, predominantly twelfth grade, from small classes in rural schools. Generally, students who elect to take physics courses have above average ability (i.e., restricted range of ability).
8. The four treatments in the experiment were standard MBL, delayed MBL, pencil-and-paper (a reference activity), and test only (no activity). Treatments were compared by analysis of covariance.

Several assumptions influence the interpretation of results.

1. Treatment groups were equivalent. Students could be randomly assigned to treatment groups only within each class. This assumption was statistically tested for the available covariates and measures of pretest performance.
2. Prior relevant knowledge and experience was either equivalent for

all students, was accounted for by covariates, or was randomly distributed among treatments. All students had been taught kinematics earlier in the year.

3. The group size did not influence learning during the treatment activity. Groups consisted of two or three students in MBL activities, and two to four students in the pencil-and-paper activity. This assumption may not be valid, given the importance of the quality of group dynamics on individual learning, which is discussed in Chapter III (unit of statistical analysis).
4. Students who were absent on any one day were equivalent to the rest of the students in the study (in ability, motivation, etc.). These students were automatically assigned to the test-only treatment.

Structure of the Dissertation

Relevant literature will be reviewed in the next chapter (Chapter II) to provide more detailed explanation for the rationale, some guidelines for designing the research, and some theoretical framework for interpreting the experimental results. There is surprisingly little known about the quality of graphing skills among school students, or the types of problems they have with graph comprehension. The research design, treatments, instruments for measuring performance, and the methods used to analyze the results are detailed in Chapter III.

Research results are examined in three chapters. Chapter IV addresses the question of treatment effects in posttest scores, and then

examines students' responses to individual items to look for what students learned--concepts or graphing skills? Chapter V analyzes results from tests of interpreting and constructing graphs, searching for information about what kinds of problems students have with graphing. Chapter VI addresses the issue of whether learning was mediated by cognitive or motivational behavior. Data in this chapter derive from questionnaires and behavioral observations. Finally, Chapter VII provides a brief summary of research findings, some considerations for implementing MBL into the classroom, and some ideas for future research.

CHAPTER II

SYNTHESIS AND INTERPRETATION OF LITERATURE

Introduction

This literature review focuses on educational and psychological theories, empirical research results, and pragmatic considerations that are relevant in formulating the research hypotheses, the research design, and the explanation of experimental results and observations. It first examines the role of science laboratories in science education, before concentrating on the specific contribution offered by microcomputer-based laboratories (MBLs). The real-time graphing feature is singled out for additional attention because it is central to the research hypotheses. The literature on cognitive constraints in information processing and recent research on intrinsic motivation of educational software are particularly relevant.

Although there is abundant evidence of conceptual difficulties that students experience in learning basic concepts in kinematics, there is neither a consensus nor a clear picture of the conceptual difficulties they seem to have in using graphs. A recently formulated theory of graph comprehension Pinker (1986) provides a framework for considering the attributes of both graph and graph user that affect graph comprehension. Particular attention will be given to evidence for

the existence of problems with graph skills and to the implications for instruction.

Laboratories in Science Education

Among the goals of laboratory experiences in science education are learning and understanding concepts, and developing logical, inquiry, and problem-solving skills. Other goals, such as promoting positive attitudes and social skills, will not be discussed here.

There seems to be some validity in the old aphorism, "Hear and forget, see and remember, do and understand." For example, laboratory activities provide opportunities for students to (a) engage in processes of investigation and inquiry (Tamir, 1977); (b) use personal experience to acquire knowledge and construct concepts (cf. Piaget's "constructivism", Ophardt, 1978); (c) exchange points of view with lab partners, thus encouraging reflection, objective verification, and self-correction (cf. Piaget's "constructivism", Kamii, 1984); and (d) acquire information via multiple senses, an approach that appears to enhance memory of specific factual experience (Esler, 1982). Hegarty (1982) suggests that one of the most powerful uses of laboratories is in removing or correcting students' misconceptions in physics. The misconception line of research is discussed later in this chapter.

Some type of structure is important in laboratory activities, particularly for students who cannot apply formal operations in their understanding of science processes (Spears & Zollman, 1977; Tobin &

Capie, 1982a), and also within self-instructional programs (Tanner, 1969). A number of studies (Charen, 1966; Coulter, 1966; Egelston, 1973; Lerch, 1973, to name just a few) have compared the influence of inductive and deductive laboratory activities at high school and college levels on various measures of performance and affect. Most of these indicate that, where there are significant differences in outcome measures, these differences favor inductive laboratories.

Although the research indicates that laboratory experiences are potentially powerful instructional tools for achieving the goals described above (e.g., Bredderman 1982, 1984; Hofstein & Lunetta, 1982; Johnson, Ryan, & Schroeder, 1974), this potential is not always fulfilled (Reif & St. John, 1979). At higher levels of education, laboratories are less investigative and more confirmatory (Hegarty, 1982; Tamir, 1977), failing to provide genuine problem-solving experiences (Charen, 1966). This is in part due to a number of pragmatic constraints, including time restrictions and the pressure to cover content. Science laboratories are expensive in terms of space, equipment, finance, and time. There are concerns about whether they are cost effective or efficient in achieving the goals of science education. The need for them should be subject to critical scrutiny (Bradley, 1968; Hofstein & Lunetta, 1982). Ways of reducing the costs should be examined, and alternatives should be explored (Menis, 1982).

Microcomputer-Based Laboratories

In MBL labs, the computer is used as a sophisticated laboratory instrument, for any or all of the following functions: (a) detecting and/or measuring physical phenomena, (b) performing calculations with the data, (c) displaying the data in any format (e.g., tables, frequency distributions, or graphs), (d) printing the data, (e) storing the data, or (f) comparing it with some other reference data. Many different kinds of physical phenomena can be detected and measured by appropriate sensors coupled via the necessary interface to the computer. If the input data are in analog format, they must be converted to digital information that the computer can handle. This is done by an analog to digital converter (the interface).

For a long time all branches of science have used equipment either run by computers or with the equivalent of an on-board computer for the detection, measurement, and manipulation of information (e.g., specific ion electrodes, automatic titrators, mass spectrophotometers). Such equipment is used extensively in research and in application within the science-based professions. Only recently have these kinds of facilities been made available for teaching labs in schools or in universities at the undergraduate level.

MBL units used in school science laboratories may be as varied as traditional labs in terms of exactly what they are trying to accomplish and how they are intended to achieve this (i.e., philosophical and psychological bases). For instance, some developers of MBL have focused on transducers, developing equipment and sensors to measure properties

needed for an experiment (e.g., Rafert & Nicklin, 1982, 1984); others have focused on computer programming as the means of adapting a given tool to serve a specified need (e.g., McInerney, 1984); still others have been content to treat both the hardware and software as a "black box" and concentrate on conceptual interpretation of the output of information (e.g., Tinker, 1981; Tinker & Stringer, 1978). This discussion will be restricted to the latter implementation, which was used in this research.

There are a number of practical considerations to the MBL units that are now available that make them an attractive instructional tool.

1. The cost has been dramatically reduced for micro-computers and peripherals such as MBLs over recent years.
2. Extensive field testing (Barclay, 1986; Linn et al., in press; Mokros, 1986; Mokros & Tinker, in press; Thornton, 1986) has helped to locate and fix technical problems, and to improve capability, flexibility, screen design, and user friendliness.
3. No prerequisite knowledge of electronics is required.
4. Flexibility and versatility of hardware (and software) is increasing with the development of an increasing number of sensors, and low-cost, general-purpose interface boxes.
5. Users do not need to be competent programmers.
6. More software, and more powerful software, is becoming available.
7. More instructional support materials are becoming available.
8. Various workshops and symposia are introducing the units and providing experience and training in their use for developers, researchers, teachers, and administrators.

Two studies (Linn et al., in press; Mokros & Tinker, in press) have found that extended experience (several weeks) with MBL activities improved middle-school students' performance in interpreting graphs, including graphs of concepts not included in the treatment period (Linn et al., in press). However, a hard core of students (about one third) continued to make the same fundamental mistakes that they had made before the treatment.

This dissertation research was designed to complement these studies, by examining in finer detail some of the characteristics of the MBL unit and of the students that impact on students learning to link a physical event with a graph, and that may explain which students benefit most from MBL experience and why some students fail to do so. Specifically, the research focused on real-time graphing as the most salient feature of MBL, and on general ability, specific graphing ability, and sex as potentially important attributes of the students.

Comparison of MBL and Conventional Laboratories

Measurement. Using the computer for detecting and measuring physical properties has a number of practical advantages over traditional (non-MBL) measuring techniques: (a) the data are generally better in quantity and quality because the computer is both faster and more accurate than students in taking measurements, (b) the time scales can be either shorter (down to a split second) or longer than can normally be accommodated in a conventional lab, and (c) properties can

be easily measured that are otherwise difficult or impractical to measure (e.g., force, motion).

Calculation. Thornton (1986) claims that using the computer for calculating and manipulating the data relieves students from tedious procedures that consume their time, attention, and mental energy, and thus allows them more time to attend to the science behind the activity. Counter arguments claim that this use of the computer will induce a "black box" problem, where students learn to manipulate the system, but lose sight of the laws under investigation (McInerney, 1984). Barclay (1986), who has worked with a wide range of activities with students from sixth grade to college, states that, in practice, these "black box" fears are not realized. Such fears seem to be more relevant to computer simulations, where the data are generated by the computer ("black box"), than to MBL activities, where the data are generated from a student-controlled event external to the computer.

Representation (tables or graphs). Using the computer for constructing graphs has similar claims of advantages, and counterclaims of "black box" disadvantages. However, MBL students do not forfeit control or responsibility for the graphing. They are able to select which parameter to graph (in the motion unit), the type of graph display (in reaction-time unit), or the scales of the axes. After the data have been collected, the students can alter any of these parameters or compare the graph with another graph stored in memory. In some instances (e.g., reaction time), students are able to remove specific

datum points from the data set and examine how this influences the represented information (including the mean value).

These three activities (measuring, calculating, and graphing) occupy most of the time available in traditional laboratories. In MBLs, students do not have to spend their time on such procedural activities but spend it instead in interpreting data and investigating phenomena by reiteration and altering variables. Because of the considerable impact both on what can be done in a lab and on how lab time is spent, MBLs constitute as much a curriculum movement as a technological change. By providing new experiences, they induce different behaviors by the students.

Real-Time Graphing

Real-time graphing of data with the motion detector coupled to the computer is fast, and it is dynamic in the sense that the line on the graph seems to grow and move across the screen as the event progresses. Both of these features (the speed and the dynamism) make the developing graph responsive to student command (or whim) and may have a considerable impact on cognition (information processing) and/or motivation. For these reasons, real-time graphing was selected as a major focus of this research, using graphs of distance and velocity. These data are generated as a student's movements are monitored by the motion detector (described in Chapter III).

Cognition

Researchers who have worked with the motion unit (Mokros, 1986; Mokros & Tinker, in press; Thornton, 1986) suggest that the linking in time of a physical event with a simultaneous graphic representation may facilitate an equivalent linking in memory. Mokros (1986) discussed the possibility that, to the extent that this occurs, the real-time graphing may operate as a bridge to formal reasoning and development.

Information-processing theory (Bransford, 1979, chapter 2) provides a framework for considering how this facilitation might occur. Short-term, or working, memory is limited in capacity (generally recognized to be in the range of five to seven chunks of information), it is limited in retention time (15 to 20 seconds in the absence of rehearsal), and it is limited in the rate at which it can transfer information to long-term memory (5-10 seconds per chunk) (Simon, 1974). This model of information processing assumes that the initial entry and processing of information in the brain takes place in short-term memory, which has a fairly short decay rate. Real-time graphing may allow rapid cognitive linking within short-term memory (e.g., of a physical event and a point on a graph) and increase the likelihood of the linked information being transferred to long-term memory as a single unit.

If incoming information exceeds the capacity limitations described above, there will be a breakdown in understanding and learning (Bransford, 1979). Students, especially those inexperienced in a particular task, are susceptible to early cognitive overload (i.e., the comprehensiveness of the chunks they can massage is restricted). In

addition, they are unsure to what features they should attend. Because movement in a display dominates attention (Bertin, 1983), real-time graphing may encourage students to selectively attend to salient points on the graph (e.g., inflection points, intersections) that correspond to changes in the physical event (e.g., changes in speed or direction). The dynamic display may also stimulate students to find out how to create changes in the graph by doing something in the physical world.

In this research, the delay in the display of the data transformed the situation by separating in time the two elements that need to be linked, presenting them serially (or sequentially) rather than in parallel (or simultaneously). There are theoretical considerations which might lead us to expect that a delay in excess of the 15-20 seconds retention time in short-term memory in the absence of rehearsal would be more effective than immediate (real-time) graphing feedback.

Given this brief delay, to make an effective cognitive linkage between the two elements (the event and the graph), students would have to actively either transfer information about the physical event to long-term memory or maintain it in short-term memory until the graph had been constructed (Bransford, 1979). To maintain the information in short-term memory, students could rehearse the event mentally, or elaborate on the information, for instance, by forming predictions or expectations of what the graph will look like. Such maintenance activities impose an additional cognitive burden, but they should improve learning and retention of information (Bransford, 1979). (Similarly, the pencil-and-paper activity could also keep students

attending to salient features of the graphs and mentally rehearsing the event as they construct the graph.)

However, students in a delayed-MBL treatment may not employ these maintenance activities during the delay because either (a) they do not know such mathemagenic techniques, (b) they are not aware that these techniques are called for by the situation, (c) the additional cognitive demand of these activities may lead to their exceeding students' cognitive capacity, or (d) they may not consider the extra amount of invested mental effort worth while. Salomon (1983) considers this perception to be an important determinant of learning. For instance, when Kirk, Kauchak, & Eggen (1978) increased the cognitive level of textual cues in a graphing task (from specific to general cues), student performance on a later test decreased significantly. They attributed this to cognitive overload.

If the task does not impose too great a cognitive burden on the student, relative to the student's cognitive load-bearing capacity, or threshold (affected by factors such as ability, interest, anxiety, prior knowledge, and its overlap with the current activity), then a slight delay in presentation of the graph may be more effective than real-time graphing in promoting a cognitive linkage between an event and a graph. The key question then, of whether the immediate graphing or the slightly delayed graphing would result in superior learning, may well hinge on whether or not the students use the delay for maintenance or elaborative rehearsal processes, that is, whether the students both know how and are motivated to engage in rehearsal activities during the delay. In this study, students were not coached in appropriate mathemagenic techniques.

Motivation

In the last few years, the availability of computer games and the need to assess the instructional effectiveness of educational software have stimulated a resurgence of interest in intrinsic motivation (Malone, 1981; White, 1984). Lepper (1985) reviewed the literature on three different theoretical traditions that have contributed to our understanding of motivation. The first group of theorists has discussed motivation principally in terms of competence, effectance, mastery, or challenge (Frase, 1968; Harter, 1978, 1981; Kagan, 1972; White, 1959). In doing so, they have focused on the role of people as problem solvers. The second group has characterized motivation in terms of perceived control and self-determination (Condry, 1977; Condry & Chambers, 1978; Deci & Porac, 1978; Rowe 1974a, 1974b). They are concerned with people as "actors who seek to exercise and validate a sense of control over the external environment" (Lepper, 1985, p. 5). The third group focuses on motivation through curiosity, incongruity, discrepancy, and complexity (Berlyne, 1957, 1966, 1978; Kagan, 1972), viewing humans as information processors. Most of the experiential science programs of the 1960s and 1970s were built on this conception of motivation. As Lepper points out, these approaches are neither distinct nor incompatible. They do, however, provide a framework for examining the behavior of students for indications of motivation, and for differences in motivation.

Guided by these diverse theories, Malone (1981) described a number of features of highly motivating computer games that appear to make the games fun, challenging, and conducive to a sense of competence and

mastery. These features often depend on being uncertain of achieving a given goal, incorporating variable difficulty levels, having multiple level goals, and presumably allowing multiple ways of achieving a given goal. Cognitive curiosity can be aroused by making students believe that their understanding or knowledge is incomplete or inconsistent. The MBL unit used in this study incorporates all of these features.

Lepper (1985) discussed several ways in which motivation and affect may interact with cognition. Motivation could influence (a) the direction and intensity of attention (Simon, 1967), (b) the depth of involvement, (c) the level of arousal, and (d) the representation of abstract problems by prompting effective means. It also could alter the perception of the mental effort invested in a task (Salomon, 1983).

Although these ideas, theories, and conceptual frameworks provide guidelines for observing and interpreting student behavior during engagement with the MBL activities, they do not provide a basis for hypothesizing which of real-time or delayed graphing conditions should be more intrinsically motivating. Again, the issue of whether an activity is challenging, encourages competence, and incites curiosity may depend on the balance between cognitive load of the task and the cognitive capacity of the user, rather then depending solely on the intrinsic attributes of the activity. We may speculate, however, that the real-time condition comes closer to the immediacy exhibited by arcade games, and thus may have more attention-focusing power.

Concepts and Misconcepts of Distance and Velocity

MBL units are available for many science courses--chemistry, biology, physiology, and physics. There are a number of reasons for selecting the motion detector and high school physics students. Within physics education, a considerable amount of research has concentrated on understanding the misconceptions students seem to have, particularly with concepts in mechanics. Research has focused on (a) characterizing the misconceptions as adhering to a consistent theoretical base (e.g., Aristotelian, Impetus, or Newtonian theories), (b) understanding their origins from everyday experience, and (c) determining how they should be remediated.

In mechanics, faulty beliefs are common and have generally been characterized as being closer to Aristotelian beliefs (Champagne, Klopfer, & Anderson, 1980; DiSessa, 1982; Whitaker, 1983), or medieval Impetus theory (Halloun & Hestenes, 1985a, 1985b; McCloskey et al., 1980), than to Newtonian beliefs. In the more specific area of kinematics (the study of motion without regard to the forces producing it), the main difficulty in understanding velocity, exhibited by college students as well as middle-school students, seems to reside in confusing distance with velocity (McDermott, 1982; Metz, 1982; Minstrell, 1982b; Trowbridge & McDermott, 1980a, 1980b).

These conceptual problems are "mis"conceptions only in the sense that they do not conform to the prevailing paradigm. Alternative names in the literature, for instance, preconceptions, preparadigms (Champagne et al., 1980), zero-order models (Clement, 1982), naive beliefs

(McCloskey, Caramazza, & Green, 1980), intuitive beliefs (Leboutet-Barrell, 1976), protoconceptions (Trowbridge & McDermott, 1980a), alternative conceptions (Minstrell, 1982b), or thinconceptions (Berger, personal communication, July, 1986) are used to convey the sense that these ideas or beliefs have emerged from the students' individual experiences in the real-world, but they are seldom the result of careful and rational consideration, being neither explicit nor integrated within a consistent belief system (McCloskey et al., 1980).

Whatever the apparent theoretical base and however the intuitive conceptions arise, researchers have found consistently that they are difficult to detect and remarkably resistant to remediation. Such conceptions are elusive. For instance, when students have been asked to explain their answers to given questions, the reasons they have provided have generally been diverse and often inconsistent (Whitaker, 1983; White, 1984). Alternative conceptions are also well camouflaged. They often go unnoticed because students' superficial knowledge of scientific jargon, formulae, and mathematical techniques can mask their problems with the underlying concepts (Clement, 1982; Halloun & Hestenes, 1985a, Leboutet-Barrell, 1976; Trowbridge & McDermott, 1980a).

Because these misconceptions have been constructed from personal experience and they are reasonably functional in everyday life, they are stable and resistant to remediation.

> These preconceptions are amazingly tenacious and resistant to extinction because of the influence of such factors as primacy and frequency, because they are typically anchored to highly stable related and antecedent preconceptions of a more inclusive nature; because they are inherently more stable . . . and lastly because resistance to the acceptance to new ideas contrary to prevailing beliefs seems to be characteristic of human learning. (Ausubel, 1968, p. 336)

When students are confronted with concepts that are inconsistent with their preconceptions (e.g., in Newtonian physics), often they either (a) fail to see the contradictions, (b) distort the Newtonian concepts to fit their existing beliefs, or (c) memorize the new concepts as separate, unconnected entities (Clement, 1982; Halloun & Hestenes, 1985b). In research designed to evaluate conceptual understanding, it is therefore necessary to state problems or questions in a format as free from potential disguises, such as scientific jargon or mathematical representation, as possible.

Teachers need to help students make explicit connections among physical concepts, their mathematical representations, their graphic representation, and their appearence in the real world, but they are unlikely to change students' beliefs just by telling them (McDermott, 1982; Minstrell, 1982a, 1982b). Because students are so convinced by their alternative conceptions and are so skillful in rationalizing them, conventional instruction seems to be not only surprisingly ineffectual in changing them (Clement, 1982; Halloun & Hestenes, 1985a; McCloskey et al., 1980), but may instead provide students with new terminology (e.g., "momentum") for expressing their erroneous beliefs (Leboutet-Barrell, 1976; McCloskey et al., 1980).

Instruction which promotes conceptual conflict appears to be more effective (Champagne et al., 1980; Minstrell, 1982a), although it may not be sufficient. Halloun and Hestenes (1985b) found that students changed their beliefs only when they realized there were inconsistencies in their own thinking. This is consistent with the curiosity theory of motivation arising from conceptual incongruity, dissonance, or

discrepancy (Berlyne, 1957, 1966, 1978; Kagan, 1972). Berlyne (1978) has defined curiosity as "an internal state occasioned when subjective uncertainty generates a tendency to engage in exploratory behavior aimed at resolving or mitigating the uncertainty" (p. 97).

Graphs, Graphing, and Graphers

Students need graphing skills because they are so powerful as logical reasoning tools for describing data, exposing relationships, and communicating results (Bertin, 1983; Bestgen, 1980; Kosslyn, 1985). Graphs can provide deep insight into the structure of information by exposing patterns, trends, relationships, and comparisons (Lefferts, 1981), thus enabling us to explore data thoroughly, to look for patterns and relationships, to confirm or disprove expected behavior, and to discover new phenomena. Macdonald-Ross (1977) summed up these uses of graphs by describing them as "visual arguments." Chambers, Cleveland, Kleiner, and Tukey (1983) went so far as to claim that no single statistical tool is as powerful for facilitating pattern recognition in a complex data array as a well-chosen graph.

Graphing is an important basic-process skill. In every discipline in the social, biological, and physical sciences; in all applied science and social and governmental policy; and in magazines and newspapers, we depend upon the appropriate use of graphs to help us interpret quantitative information (Macdonald-Ross, 1977). Two recent additions to the market place of job skills have provided a new impetus of

interest in these skills. New statistical analysis techniques make extensive use of graphical analysis (Anscombe, 1973; Chambers et al., 1983, Cox, 1978; Fienberg, 1979; Wainer, 1974). Revolutionary improvements in the capability and availabilty of computer graphics have created huge demands for people who can use these capabilities competently and effectively (McLellan, 1980).

Graphing is taught in all the major science programs at the elementary level whose development was supported by National Science Foundation (e.g., SAPA--Science: A Process Approach, SCIS--Science Curriculum Study, ESS--Elementary Science Study, COPES--Conceptually-Oriented Processes in Elementary Science). Items to test graphing skills are included in science process instruments (e.g., Dillashaw & Okey, 1980; Doran, 1978; Fraser, 1980; Padilla, Okey, & Dillashaw, 1983; Tannenbaum, 1971; Tobin & Capie, 1982b). Graphs are used as an important method of presenting information in secondary science textbooks. By the time students enter college, they are assumed to be competent in graphing skills, among other science process skills (Arons, 1979).

The literature contains diverse definitions and uses of the term "graph." At one extreme, Fry's (1981) definition is synonymous with general usage of "graphic" and is really too broad to be useful for research purposes. At the other extreme, Macdonald-Ross' (1977) definition is effectively restricted to Cartesian graphs, and may be too narrow to be very useful for educational purposes. Somewhere in between these, Pinker (1986) states the definition used in this dissertation.

> All graphs try to communicate to the reader a set of pairings of values on two or more mathematical scales, using depicted objects

> whose visual dimensions . . . correspond to the respective mathematical scales, and whose values on each dimension . . . are proportional to the values on the corresponding scales. (p. 2)

In effect, graphs have many of the properties of verbal language in providing a common system for representing and exchanging information. Encoding information into a graph (graph construction) and decoding information from a graph (graph comprehension) rely on the understanding and strict application of rules of grammar, syntax, and convention. These constraints impose order and meaning on the assembled graph. Compared with verbal language, graphs are an information-dense system, being very effective in distilling a lot of information into a small amount of space. The arrangement of this information in the graph must be unambiguous to prevent misunderstanding between the graph constructor and the graph interpreter.

Graph Construction

Macdonald-Ross (1977) and Pinker (1986) have reviewed results from a number of empirical studies comparing the efficiency of different ways of representing data--text, tables, charts, and graphs. Results generally indicate that (a) tables are best for illustrating absolute values, (b) bar graphs (bar charts) are best for comparing differences among dependent variable values for specific independent variable values, (c) pie graphs (pie charts) are best for comparing proportions, and (d) line graphs are best for illustrating trends and interactions.

Once we have decided to use a graph in a given context, we need to know what features contribute to making it efficient and effective.

Learning from graphs appears to be a function of the form in which the data are presented, that is, (a) the quantity of data (Bertin, 1983; Price, Martuza, & Crouse, 1974; Washburne, 1927), (b) the complexity or abstractness of the information (Chambers et al., 1983; Macdonald-Ross, 1977), and (c) both the visual (Cleveland, 1984b; Cleveland & McGill, 1984) and logical (Washburne, 1927) organization of the data.

Much of the designer's art goes to make the information more immediate and to reduce confusion. "Simplification is an obligation of the communication process" (Bertin, 1983, p. 166). In addition to capitalizing on perceptual mechanisms of the reader, the designer conforms to and utilizes the visual grammar and syntax that are components of the graph schema. Guidelines for the construction of graphs are available in several "how-to" texts and manuals (e.g., Cleveland, 1985; Cleveland & McGill, 1984; Guidry & Frye, 1968; Lefferts, 1981; Schmid, 1983; Tufte, 1983). The essence of these guidelines is expressed succinctly by Cleveland (1985) as clear vision and clear understanding.

Tufte (1983) describes the same thing eloquently in his principles of graphical excellence:

> Graphical excellence is the well-designed presentation of interesting data--a matter of <u>substance</u>, of <u>statistics</u>, and of <u>design</u>.
>
> Graphical excellence consists of complex ideas communicated with clarity, precision, and efficiency.
>
> Graphical excellence is that which gives to the viewer the greatest number of ideas in the shortest time with the least ink in the smallest space.
>
> Graphical excellence is nearly always multivariate.
>
> And graphical excellence requires telling the truth about the data. (p. 51)

Comprehension of Graphs

Graphs are effective because they make optimal use of general cognitive and perceptual processes. The eye-brain system can summarize vast amounts of information quickly and extract salient features, but at the same time it is capable of focusing on detail (Chambers et al., 1983). In other words, graphs allow us to see the leaves, the branches, and the whole tree (Bertin, 1983).

Reading and interpreting graphs involves two processes--visual perception and graphical cognition (Cohen, 1981; Moyer & Bayer, 1976; Pinker, 1986; Reicher, 1969). Visual perception is the process of detecting the separate features of the visual display, and deriving visual (or semantic) meaning from the spatial organization of the elements. Pinker (1986) has recently developed a theory of graph comprehension (conceptual framework shown in Figure 2-1), synthesizing ideas from such diverse sources as Bertin's (1981, 1983) seminal work on graphs, charts, and maps; Cleveland and McGill's work (Cleveland, 1984a, 1984b, 1985; Cleveland & McGill, 1984, 1985) on visual perception and visual constraints; Kosslyn's (1985) ideas on constraints on processing information from graphs; research on cognitive representation of visual information (Marr, 1982); and diverse research on psychophysics and psychology.

Graphical cognition relies on the existence or the development of an appropriate graph schema. A schema is a representation in memory that incorporates knowledge of a domain (Bransford, 1979, chapter 6; Pinker, 1986). This knowledge includes both a semantic description,

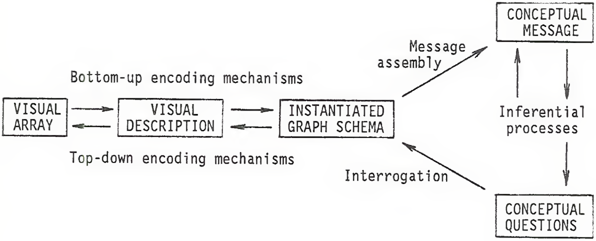


FIGURE 2-1. Framework for a theory of graph comprehension (reproduced with permission from Pinker, 1986, Fig. 19).

with unquantified variables defined, and a syntax of relationships among the variables. According to Pinker, a general graph schema embodies knowledge of what graphs are for and how they are interpreted in general. Such knowledge includes an understanding of the following: (a) a one-to-one correspondence between a real-world variable and its representation by a visual attribute on a graph (Shepard, 1975); (b) a more abstract relationship between the structural functions of an external object and the functional relations of the graphic representation (Shepard, 1975); (c) the ratio magnitudes of attributes are specified in terms of the coordinate system; and (d) textual materials (labels, scales, etc.) specify pairings of the values of the attribute with the coordinate system and the real-world event (Pinker, 1986). In tasks requiring graph comprehension, the graph schema

specifies how to (a) translate information in a visual description into the conceptual message, (b) translate the conceptual question into a search strategy, and (c) recognize which type of graph, complete with pertinent grammar, is appropriate.

Graphical cognition is the process of extracting conceptual information from the visual image in short-term memory according to (a) the syntax of the appropriate graph schema, (b) selective attention necessitated by the limited capacity of working memory, and (c) the salience of the various elements of the image (Pinker, 1986). Salience, or the likelihood of a given element being encoded, will be influenced by innate attributes of the graph (e.g., dynamic display, color, shape, etc.), and by recency and primacy effects (Bransford, 1979, p. 171; Pinker, 1986).

There seem to be two fundamentally different ways of extracting information from a variety of contexts (Schneider & Shiffrin, 1977; Shiffrin & Schneider, 1977), including graphs (Pinker, 1986)--either by recognition (bottom-up processing) or by searching (top-down processing). Thus Bertin (1983) distinguished between "seeing" graphs and asking questions of them.

In bottom-up encoding, elements in long-term memory appear to be automatically activated by the visual image itself, proceeding without subject control, without stressing capacity limitations, without necessarily demanding attention, and without interfering with other ongoing mental processes (Posner & Snyder, 1975; Schneider & Shiffrin, 1977). This kind of automatic message assembly develops following consistent mapping of stimuli to responses over a number of trials.

Practice and reiteration provided in MBL activities should promote an automatic linkage between the event and the graph, thus reducing the cognitive demand of interpreting the graph.

Top-down encoding (or interrogation, or controlled search) retrieves or encodes new information on the basis of conceptual questions (i.e., information that the reader wants to extract from the context). For instance, hypotheses, textual cues, and contextual purpose induce searching behaviors. This appears to require attention, be capacity-limited (Posner & Snyder, 1975), and be done serially (Reicher, 1969; Schneider & Shiffrin, 1977).

Many cognitive tasks may be viewed as combining automatic activation and conscious strategies. For instance, the search process generates an enriched, or specialized conceptual message. Applying logical or mathematical inference rules to this conceptual message can, in turn, generate new conceptual questions (Pinker, 1986). MBL activities provide good examples of this. They combine observation (bottom-up encoding) with predicting outcomes, synthesizing information about specific instances into laws or concepts, and testing hypotheses (top-down encoding of the conceptual message).

In theory, any quantitative information that is resident in the graph can be extracted from it. In practice, some information is more easily extracted than other information. Because of the limitations of time and memory, Pinker (1986) suggests that comprehension will be increasingly difficult as it requires recognition, search, and inference processes to extract the necessary information (i.e., difficulty increases with cognitive level of the comprehension task). Although an

increasing degree of abstractness or complexity makes graphs more difficult to comprehend, it also makes graphs more efficient than alternative systems of representation for clear communication.

Problems in Graph Comprehension

A graph is successful only if the decoding is effective (Cleveland & McGill, 1985). Individuals may have difficulty with comprehension if either (a) their graph schema is inadequate, or (b) the graph is inadequate (i.e., the necessary elements of the graph are either not easily perceived or not salient, or the rules of syntax are flouted) (Pinker, 1986). Although practice can improve the probability of encoding information and of developing associative links (Linn et al., in press; Mokros & Tinker, in press), it may be more difficult to improve students' graph schemas.

Students are taught how to read only certain types of graphs and we expect them to be able to generalize from these to the many forms that are created. Pinker (1986) suggests that there are three ways for students to enrich their graph schemas with appropriate links, each requiring increasing insight on the part of the student: (a) they can be told, as in formal, didactic instruction; (b) they can generalize from instances by induction; and (c) they can infer, or deduce, from their prior knowledge (e.g., from knowledge of the behavior of the property being graphed, or of mathematical relations).

In the research described in this dissertation, the MBL activities should have provided opportunities for both inductive and deductive

enrichment of students' understanding about the general properties of line graphs. In addition, theory (e.g., Piaget's ideas on constructivism discussed by Kamii, 1984) and empirical evidence with sentence construction (Bobrow & Bower, 1969) suggest that students who work with data that they generate for themselves are more likely to recall linkages and associations in the data.

Measuring Graphing Skills

It is difficult to assess the status of graphing skills because of the lack of appropriate instruments and the paucity of theoretical bases on which to construct such tests. We need tests for measuring components of graphing ability, and test items need to be examined carefully for construct validity (Price et al., 1974).

A test of graphing skills has been developed by McKenzie and Padilla (1986) for use with students from grade 7-12. This test has two major sections: construction of graphs and interpretation of graphs, that are further broken down into specific skills. The structure of the test items and and the performance of students (grades 7-12) on each of these skills is shown in Table 2-1 (Padilla, McKenzie, & Shaw, 1985). The structure of this test guided the design of the graph construction exercise and the evaluation of student performance in this research.

Most authors who have studied graphing performance agree that students' ability to read graphs is superficial. As shown in Table 2-1, students seem to perform quite well as long as they are being tested on tasks involving direct reading of specific data from tables or graphs,

Table 2-1. Performance of students from grades 7-12 on each category of the Test of Graphing Skills (Padilla et al., 1985).

Graphing skill		(% maximum score)
Construction	scaling axes	32
	assigning variables to axes	46
	plotting points	84
	using best fit line	26
Interpretation	determining coordinates	84
	interpolate/extrapolate	57
	describe relationship	49
	interrelating graphs	47

but their performance declines considerably on tasks that require interpretation (Bestgen, 1980; Fry, 1981; Padilla, McKenzie, & Shaw, 1986; Vernon, 1946, 1950). This consensus implies that students develop operational algorithms for working with graphs rather than concepts or executive strategies.

Student Characteristics and Graphing Skills

Surprisingly little is known about the influence of student attributes on the comprehension of information from graphs. Predictably, several studies show graphing skills are correlated with various measures of general intelligence (Culbertson & Powers, 1959), ability (Kauchak, Eggen, & Kirk, 1978; Kellogg, 1967), or reasoning and development (McKenzie & Padilla, 1984). Several studies have shown an

increase in performance with age or grade level of students (Fraser, 1980; Padilla et al., 1983; 1985), which probably translates to increased experience and accompanying development, both of which are related to graphing abilities. Experience with graphing activities improves graphing skills, whether the activities are laboratory-discovery or demonstration-discussion (Kellogg, 1967), hands-on or simulation (McKenzie & Padilla, 1984), or MBLs (Linn et al., in press; Mokros, 1986; Mokros & Tinker, in press).

An examination of the graphing skills for differential gender effects was an important component of this study because of (a) the historical cultural bias against females taking courses in the physical sciences, (b) potential differences in spatial skills, and (c) the correlation between spatial visualization and performance of high school and college students in mathematics, science, engineering, and chemistry (Baker & Talley, 1972, 1974; Fennema & Sherman, 1977; Sherman, 1980). Although there is a considerable body of research dealing with gender differences in spatial visualization ability, there do not appear to be any data relating graphing ability to either spatial visualization ability or gender.

Males generally have better spatial skills than females (Coltheart, Hull, & Slater, 1975; Eliot & Fralley, 1976; Sherman, 1967, 1980), but frequently these differences are not apparent until adolescence (Sherman, 1980; Smith & Schroeder, 1979). This has produced vigorous debate as to whether this is a fault of the measurement instruments (Eliot & Fralley, 1976), and whether it results from the accumulated cultural differentiation between the sexes in experiences that develop

spatial skills (Hilton, 1985; Sherman, 1967) or is predominantly genetically determined (Garron, 1970; Hartlage, 1970; Stafford, 1961).

Graphing Misconceptions

Students appear to hold fundamental misconceptions about graphs. Clement, Mokros, & Schultz (1986) and Mokros and Tinker (in press) have studied specific kinds of problems exhibited by middle-school students during tasks of graphing comprehension. The most frequent seemed to be confusion between the slope and height of points, and the tendency to see the graph as a picture rather than a symbolic representation of information. Vernon (1946) commented that adults had the same difficulty with complex graphing tasks. They sometimes visualized the charts clearly without translating the images into the information the chart was intended to convey.

Mokros and Tinker (in press), claiming that these difficulties seemed to be readily removed by MBL experiences, questioned the appropriateness of labelling them "misconceptions." To qualify this comment, it is worth noting that their own data indicated that, after a minimum of 20 classes with MBL, the "graph-as-picture" problem, and the "slope-height" confusions each persisted for about one third of the middle-school students. That certainly indicates some kind of resistance. Similar findings appeared with eighth grade students who carried out 54 MBL activities investigating temperature over a 13-week period, and an additional 25 activities in a 5-week period to investigate chemistry concepts (Linn et al., in press). Students showed

substantial improvement in their performance in identifying graph features and templates.

In such tests of graphing skills, there are two factors that are woven together--the ability to read or translate graphs *per se*, and knowledge of the conceptual content of the graph. These abilities may be independent but are not distinguished in tests of interpretation of science graphs. For instance, the slope/height confusion is usually reported in connection with distance and velocity graphs, but it is difficult to separate the slope/height confusion (problems with the system of representation) from the distance/velocity misconceptions (problems with the concepts) which have been discussed by McDermott (1982). Similarly, it is not clear whether the errors that remained after extended MBL experience in the two studies described above (Linn et al., in press; Mokros & Tinker, in press) are the result of conceptual problems (the content) or representational problems (the graph format). Because of the evidence, cited earlier, that the physics concepts are particularly resistant to change, it seems likely that the remaining errors indicate conceptual problems.

Graphs in Instructional Context

In addition to reading meaning into the relevant graphic symbols, and integrating the symbols with each other, graph comprehension requires that the graphic material be integrated with the support materials (Malter, 1948). Graphs have different degrees of contextual embeddedness--within text, within subject matter, and within an

experience or situation. Text-plus-graph is the usual teaching and communicating apparatus. Macdonald-Ross (1977) suggests that this format works best when the text does not just repeat points made by the graph but directs, comments, explains, and questions (i.e., initiates top-down encoding behaviors).

Tasks requiring application and interpretation are more embedded in context, so it would be reasonable to expect contextual experience to provide cues to facilitate performance in remembering or interpreting graphs. Most lessons on graphing occur within mathematics courses where the data in the graph are not directly related to students' activities. Many people believe that children would understand graphs better if they collect the data themselves (Swenson, 1973, cited in Slaughter, 1983; Mokros & Tinker, in press), but research fails to support this contention. Neither Kellogg (1966) nor McKenzie and Padilla (1984) found significant differences in graphing performance between subjects who collected their own data and those who did not (simulation or demonstration-discussion). Such null findings may relate to the cognitive level of the test task.

Leonard and Lowery (1984) have reviewed the available literature on supplementary cues or adjunct questions in textual material. Although there are a few conflicting results, in general (a) such questions improve learning performance, (b) review questions are more effective than preview questions, and (c) higher level questions are more effective than recall and factual questions. Similar results have been found on worksheets, where questions were positioned after presentation of graphs (Eggen, Kauchak, & Kirk, 1978; Kauchak et al., 1978; Kirk et

al., 1978; Washburne, 1927). Supplementary cues appear to be effective because they induce searching or attending behaviors in the reader, and because they focus the reader's attention on salient aspects of the graphs.

The use of higher level cues, such as answering application questions, in textual materials "facilitates later performance by encouraging students to process the content of the instruction more thoroughly, in fact to transform it, in the effort to apply it in a new situation" (Watts & Anderson, 1971, p. 393). This facilitation effect was not found by Kauchak and coworkers (Eggen et al., 1978; Kauchak et al., 1978) in their study of comprehension of information from graphs. Contrary to their expectations, high level cues hindered rather than helped college students in recalling information from graphs, even though they had answered the cue questions adequately. The researchers suggested that high level cues may have created cognitive overload--caused by the task of moving to and from graph and cue, coupled with the uncertainty of their response. In further research (Kirk et al., 1978), they suggested that there may be a threshold effect for cognitive overload. For students who are above the threshold and can cope with the load, cues that have a greater processing demand may be more effective than low-demand cues. But for those students at or below the threshold, low-demand cues would be more effective. This again highlights the possibility that interpreting graphs may place a heavy cognitive demand on students compared with comprehending textual material.

These research findings guided the development of student worksheets for this dissertation. The worksheets include supplementary cues in the form of specific synthesis questions positioned after the activity. These findings also highlight the subtlety of matching cognitive demand of a given task with the cognitive capacity of the students.

Summary

Science laboratories have the potential for making an important contribution to science education. However, pragmatic constraints have limited the fulfillment of this potential. Any new development in this arena (MBL, simulated experiments, etc.) should be carefully evaluated in view of these constraints and the goals for science education. Because MBLs provide teachers with far greater choice in time management and educational objectives, they constitute as much a curriculum change as an instructional tool.

The real-time graphing feature of MBL is compelling. Although researchers and developers suggest that the simultaneous display of the graph with the physical event helps students to link them, the literature on information processing also provides rationale for there being cognitive advantages for a slight delay before displaying the graph. Whether these advantages are realized may depend on both the cognitive demand of the task and on motivation of the students.

Students may experience difficulty with graphs because either (a) they do not have the appropriate graph schema, (b) they do not know how to use this schema to search for and extract relevant information, or (c) their experience with graphs is so restricted that the visual perception of a graph provides little conceptual meaning. Using graphs as the medium for learning about the behavior of physical properties should provide students with practice needed to develop automatic visual perception of meaning in graphs. Such experience should encourage students to construct and enrich their graph schemas, and to appreciate some of the purpose and potential value of graphs.

The MBL motion unit was selected for this research in part because of the abundant literature atesting to both the prevalence of conceptual difficulties in this subject, and to the resistance of misconceptions to alteration during conventional instruction. Researchers suggest that this MBL unit facilitates improvements in students' understanding of both physics concepts and graphs as a representational system. The task in this study was to examine this claim, looking at specific features of the MBL unit, characteristics of the student learner, and differentiating, where possible, between improvements in conceptual understanding and representational understanding.

CHAPTER III
METHODS

Introduction

This research was driven by an interest in examining the attributes of the MBL tool and the student learners in order to determine how and why a specific MBL activity is effective in generating a cognitive link between an event and a graph. Four questions represented the focus of the research interest.

1. Is the motion microcomputer-based laboratory (MBL) activity more effective than a pencil-and-paper activity (representing manual calculation and graphing of data)
   a) for learning to interpret graphs, and
   b) for learning concepts of distance and velocity?
   These skills were measured by the ability to translate between verbal and graphic descriptions of physical motion events.
2. What changes in understanding of kinematics concepts and of graphic representation are due specifically to the immediacy of real-time graphing, as separated from other features of the unit, such as
   a) dynamic graphing of data,
   b) graphing performed by computer rather than students, or
   c) increased opportunity for practice or repetition?

3. What kinds of problems do students have--in understanding, interpreting, and constructing graphs?
4. What kinds of students have problems with graphs? After partitioning variance due to ability, attitude, sex, or other attributes, are there significant treatment effects attributable to the immediate, event-responsive graphic display of standard MBLs?

Hypotheses

The main alternative hypothesis was that real-time graphing of data about experimental events during a kinematics MBL activity, in comparison with delayed graphing (about 20 seconds) of the same data, and in comparison with a pencil-and-paper reference activity, would have a significant impact on the development of a cognitive connection between a physical phenomenon and its representation on a line graph. Such improvement was measured in this study by a reduction in the error rate and/or the performance time for translating between verbal and graphic representations of a single physical event.

The second alternative hypothesis was that, if improvement in students' performance is due to features other than real-time graphing, then that performance (on the category of tasks described above) would be better after delayed-graph experience than it would be after the pencil-and-paper treatment activity.

The third alternative hypothesis was that, if information processing during a short delay enhances learning and remembering, then

performance after delayed-graph treatment could be better than after real-time graphing. However, whether the real-time or delayed graphing proves more effective may depend on the type of information processing undertaken and the motivation of the students.

The terms of the hypotheses need to be qualified and defined. These restrictions, which limit the generalizability of the results, will be elaborated in this chapter.

1. Real-time graphing of data indicates a procedure in which data are generated from a physical event and displayed in graph format immediately and automatically as the event is occurring.
2. Delayed graphing of data indicates that data are displayed in graph format after completion of the event generating the data, a delay of about 20-30 seconds.
3. The MBL hardware and software were developed by Technical Educational Research Centers (TERC). The specific probe used in this research was the sonic rangefinder, or motion detector. The software displays data from this probe as either a distance-time or velocity-time graph.
4. The MBL activity was a single-class-period, inductive, small-group activity where both teacher and experimenter were minimally involved. Guidance was provided by worksheets.
5. The content of the laboratory unit was the kinematics concepts of distance and velocity, but not acceleration.
6. The graphs used as the medium of communicating data were line graphs in a Cartesian coordinate system.
7. The student population consisted of high school physics students,

predominantly twelth grade, from small classes in rural schools.

8. The treatments in the experiment were standard MBL, delayed MBL, pencil-and-paper (reference activity), and test only (no activity).

The four treatments allowed testing of three statistical hypotheses. Given the alternative hypotheses, directional statistical hypotheses were not appropriate.

Hypothesis 1: Students from the standard-MBL treatment will not perform significantly ($p<.05$) differently in the posttests than students in the pencil-and-paper treatment. (Is MBL effective for learning concepts, given similar exposure to the concepts?)

Hypothesis 2: Students from the standard-MBL treatment will not perform significantly ($p<.05$) differently in the posttests than students in the delayed-MBL treatment. (Is the real-time feature of the MBL responsible for producing the linkage between concepts and graphic representations of them as tested in the posttest?)

Hypothesis 3: Students from the delayed-MBL treatment will not perform significantly ($p<.05$) differently in the posttests than students in the pencil-and-paper treatment. (Are features of MBL other than real-time graphing effective?)

Experimental Design

The experiment extended over three days for each class, with one day each for pretests and orientation, treatment activity, and posttests and debriefing. Research data obtained on each day and available for analysis are summarized in Table 3-1, together with abbreviations.

Data collection extended over five weeks, from the beginning of April to the middle of May. This posed a problem for external validity, because this was very close to the end of the academic year and graduation for the seniors. At this time of the year, senior students are notorious for increasing severity of so-called "senioritis," a disinclination to take school activities seriously.

The research was designed to have one group of students for each treatment (standard MBL, delayed MBL, pencil and paper, and test only) within each class (i.e., balanced design). Within each class, students who were present on the second day of the study were randomly assigned to one of the four treatments, with the restriction of a minimum of two and a maximum of three students in each of the MBL treatments.

The MBL activities offered three separate roles for students in each group. They were the controller, who operated the computer, the coordinator, who choreographed the physical events with the time variable, and the mover, who performed the motion event. Typically, students rotated among these roles throughout the activity. Developers using this MBL motion unit found that, in a field trial with sixth grade students, groups of three (and pairs on the few occasions they occurred) functioned far more effectively than groups of four. The fourth person typically had no task, and this sometimes led to a loss of interest in the activity (Mokros, personal communication, November, 1985). This design assumes that the laboratory experience is equivalent for groups of two and three students. Group sizes of two to four students were allowed for the pencil-and-paper activity.

TABLE 3-1. Summary of measures available for data analysis--description of the instrument or measure, maximum score, when it was administered, and abbreviation used throughout the dissertation.

Description	Max. score	Abbrev.
Day 1 - Pretreatment		
Information provided by subjects		
Age (in months)		Age
Sex		Sex
Number of mathematics courses (grades 9-12)		Math
Number of science courses (grades 9-12)		Science
Scholastic Aptitude Test		SAT
American College Test		ACT
Information obtained for this research		
French Verbal V-1 Test	36	Verbal
Inventory of Piagetian Developmental Tests (subset of items)	32	DEV
Relating Graphs to Events (pretest)	20	Graph
Subtest--Miscellaneous Graphs	10	Graph-Misc
Subtest--Speed Graphs	10	Graph-Speed
Kinematics Pretest (distance and velocity)	9	Pretest
Subtest--Distance Graphs	4	Pre-D
Subtest--Velocity Graphs	5	Pre-V
Questionnaire--Attitude to graphs (6 items)		
Day 2 - Treatment		
Audio recording (MBL groups only)		
Behavioral observations (journal entries)		
Experimentally-derived artifacts (worksheet entries)		
Day 3 - Posttreatment		
Relating Graphs to Events (posttest)	20	Graph
Subtest--Miscellaneous Graphs	10	Graph-Misc
Subtest--Speed Graphs	10	Graph-Speed
Kinematics Posttest (distance and velocity)	24	Posttest
Subtest--Distance Graphs	8	Post-D
Subtest--Velocity Graphs	16	Post-V
Graph construction		
Questionnaire--Attitude to treatment activity (9 items)		

Where class sizes were too small to allow a minimum of two students in each group (i.e., seven or fewer students available in a class), students were assigned to the two MBL treatments and the pencil-and-paper treatment in preference to the test-only treatment. In one school, the absentee rate was so high (averaging 18% each day) that complete data were obtained for only one student in two groups. Despite absences during the experiment, complete data were obtained for 18-20 students (within 6-7 groups) in each of the four treatments. The numbers of such students (who provided complete experimental data) in each treatment group are provided in Table 3-2.

TABLE 3-2. Number of students in each treatment group (within each class) who provided complete experimental data, number of students who provided pretest data only (i.e., were absent for part of the study), and students from the pilot-study who were included for pretest data only.

School	Standard MBL	Delayed MBL	Pencil Paper	Test Only	Pretest Only
1[a]	-	4	-	1	1
2[a]	3	-	-	-	2
3	1[b]	2	1[b]	3	3
4	3	3	3	-	-
5	2[b]	2	2	1[b]	2
6	3	3	4	4	-
7	3	3	-	1	1
8	3	3	5	4	-
9	3	3	3	8	-

a Pilot study schools. After the pilot study, minor changes were made to the worksheets, and to the posttest. Data from these schools were not included in analysis of posttest results.

b Student absent for posttest.

Treatments

The four treatments in the experiment are described in detail later in this section. In outline, they were

1. Standard MBL used the MBL with motion detector and standard software that displays data (distance or velocity) in real time.
2. Delayed MBL used the same MBL with motion detector, but the software was altered to delay the data display until after the data had been collected. The experimental event and the data display occurred serially, or sequentially, rather than simultaneously.
3. Pencil-and-paper graph-construction activity paralleled, as far as possible, the foci of attention in the two MBL activities. Students worked with the same basic events, worksheet questions, and concepts. The main differences were that (a) they did not perform the physical events to generate the data, (b) they had to calculate the data and manually construct the graphs, and (c) there was no time available for repetition and practice.
4. Test only (no activity) was included in the experimental design to statistically account for any improvement in performance resulting from recency and primacy effects.

The treatment activities were limited to a single class period. Although this is too short a period to anticipate major learning or alterations to conceptual understanding to occur, it also has some real advantages. First, it allows us to look at the issue of how short a period can generate differences in outcome. Given the need to critically assess the time and other costs of laboratory activities

(discussed in Chapter II), this issue has important pragmatic implications. Second, it allows us to investigate the initial impact, the "first-strike" effectiveness, in a comparatively simple task that is basic to graphing ability. Finally, the shorter the time interval, the easier it is to restrict the number of experimental variables contributing to the outcome. In her extended MBL study, Jan Mokros (personal communication, May, 1986) found that it was effectively impossible to separate the effects of teacher and instructional context, from the effects of MBL as tool or laboratory facility.

Standard MBL

Any MBL laboratory tool consists of two parts: hardware (computer, interface, and appropriate sensor) and software for managing data. In this MBL unit, a motion detector is used to teach relationships between position and velocity of objects as a function of time. Hardware peripherals (sensor and interface), and software were developed by Technical Education Research Centers, Cambridge, MA. Developers from TERC have also produced a series of inductive laboratory activities which, with some changes in the accompanying written materials (worksheets), have been used at both middle-school level (Barclay, 1986; Mokros, 1986) and in two college physics courses primarily designed for humanities students (Thornton, 1986).

The sensor was developed from a sonic transducer used as a rangefinder in Polaroid cameras as described by Ciarcia (1984).

> The motion probe is essentially a SONAR unit that sends out short pulses of high frequency sound (50 kHZ), then detects and amplifies

> the echo (much as a bat does). A microcomputer can then measure the time between the transmitted and received pulse and calculate position, velocity and acceleration of the object causing the reflection. The motion detector can accurately detect objects between 0.5 and 5 meters. It detects the closest object in a roughly 30 degree cone. (Thornton, 1986, p. 2)

The software (TERC, 1985) is designed for the microcomputer and probe to automate data collection and representation. The computer program is menu-driven, so students do not have to know anything about computers or programming. The software version used in this study has fewer options and a simpler menu structure than the current version. Students could produce distance-time and velocity-time graphs, but the option of graphing acceleration was disabled for this study.

For this research, worksheets developed by Thornton (1986) for use with college humanities students were pilot-tested with three groups of students--12 tenth-grade physics students from an urban University laboratory school, 4 twelfth-grade students from an urban private school, and 5 twelfth-grade students from a rural school. With these worksheets, students copied graphs from the screen display to the worksheets, partly to get them to attend to the features of the graph and partly to get a hard copy for later reference use.

After testing with the tenth-grade group, the worksheets were modified to explicitly require more active processing of the information displayed in the graph early in the activity. In the final worksheets (Appendix A), students were asked to compare graphs of different motions and extract patterns from them. In other words, they had to synthesize specific examples of graphs into templates representing general properties of motion. According to Watts and Anderson (1971), this kind of information processing should facilitate later performance.

The activities that involved the computer were basically inductive. Concepts of graphs and concepts of motion were developed in parallel. That is, the graphs were used as the main medium for representing data and building the concepts of motion. At the same time, students' understanding of the properties of motion was harnessed in helping to build an understanding of how graphs convey information.

Students were guided through the activity by the structure (conceptual structure, not procedural structure) of the worksheets. That is, the worksheets were not a verification exercise. They provided a series of goals to accomplish, and gave some suggestions of things to attend to, but they did not specify the procedures to be used. There were three types of activities on the worksheet. The first were familiarization exercises, introducing students to the equipment and requiring them to experiment with the effect on distance and velocity graphs of altering the speed and direction of the motion. In the second phase, students were asked to predict, and draw their prediction on their worksheet, what the distance and velocity graphs would look like for a complex event--a sequence of three simple constant velocity movements. Specifically, they were asked to predict and draw on their worksheets the distance (or velocity) graphs that would result from walking slowly away from the detector, stopping for a brief interval, then walking more rapidly back to the start. Times and distances were specified. Finally, they were given a graph and challenged to reproduce it using the equipment. The graphs were complex distance and velocity graphs, composed of a sequence of five simple constant velocity movements. As the students moved in relation to the sensor in order to

reproduce these patterns, they had to pay attention to the general shape of the graph, the time scale, and the distance or velocity scale. In this phase, students were confronted by any discrepancy between their expectations (conceptions) and their observations (reality).

Some of the structure was also designed to encourage discussion among students on their understanding of the essential features of the graphs they produce. At the end of each phase, there were a few supplementary questions, most of which required a single word answer. These explicitly drew students' attention to several important and salient features of distance and velocity graphs. These key points were as follows:

1. Distance graphs are not the same as velocity graphs.
2. Direction of movement makes a difference in the slope of a distance graph and in the sign on a velocity graph.
3. When an object is moving at a steady (constant) speed, the distance graph is sloping (nonzero slope) while the velocity graph is flat (zero slope).
4. When an object reverses direction, its instantaneous velocity is zero.

This single-period activity was not intended to be a complete instructional unit. It was intended to provide a single opportunity for students to experience real-time graphing of whole-body movements, under conditions that minimize dependence on the instructional activities of the teacher or experimenter. Since the data were gathered in the spring, and kinematics is commonly taught during the fall term, all the students had had some exposure to physics content.

Delayed MBL

The basic MBL software was modified slightly so that the two MBL experimental treatments differed in only one significant respect--immediacy of display. In one treatment, data were graphed in real time; in the other treatment, the display of data was delayed. On completion of a data-gathering period (approximately 20 seconds), students simply had to press a key to display the data one point at a time, at exactly the same rate as real time. It was necessary to separate the effects of real-time from those of a dynamic display to prevent confounding influences of these two salient features (immediacy and dynamism). Real-time display of data is necessarily dynamic, but a dynamic display does not have to occur in real time.

Standard- and delayed-MBL treatments shared many features. Both provided dynamic displays of the data and allowed opportunity for practice and repetition. Students performed the same tasks, fulfilled the same roles, and used the same worksheets. The computer performed the routine jobs of gathering data; transforming these to data about distance, velocity, and time; and then displaying these. The only real difference was in the timing of the graph display.

An attempt was made to modify the delayed-MBL software to remove the dynamic feature from the display of data. Essentially, the display would be stored on the computer's second screen, and could be displayed as a completed graph on a single keystroke. This was impractical because computer memory assigned to the second screen was already used by the software for data storage, and there was not sufficient random

access memory. Additional constraints on availability of equipment and students contributed to a decision to concentrate on a single feature of MBLs (i.e., real-time graphing) in this experiment.

Pencil and Paper

The pencil-and-paper, graph-construction activity was designed to parallel, as far as possible, the MBL activities. There were several reasons, apart from practical ones, for using a pencil-and-paper activity rather than a laboratory activity, where students generate their own data. First, to cover the same range of activities, a non-MBL activity would have required a week, instead of the day with MBL. Second, the pencil-and-paper activity was designed for students to be fully engaged on graphing tasks that were directly relevant to the posttest. It was conceptually frugal, because attention was not divided among procedures--measuring, calculating, interpreting, etc. Third, it allowed students to work with as many different graphs as possible within the time limits to maximize practice effects. Fourth, it provided an appropriate control for the possibility that students may benefit more in the posttest by constructing graphs than by using them.

In addition to these considerations, the worksheets (Appendix B) were designed to maximize similarities in information processing between the MBL experimental and pencil-and-paper treatments. The motion events that these students graphed were verbal descriptions of the same events as those performed by the students in the second and third phases of the

MBL treatments. Most of the supplementary cues on their worksheets were the same as, or equivalent to, those on the MBL-treatment worksheets.

In the pencil-and-paper treatment, students constructed distance-time and velocity-time graphs for two complex motion events, composed of a sequence of movements at different constant velocities. Finally they were given a speed graph of someone driving a car, and they were asked to write a story describing the events of the journey, explaining certain features of the graph, and estimating the distance of the trip. This part of the activity was adapted from one described by Woodward and Byrd (1984). Students worked in groups and were encouraged to work together, and to discuss their graphing.

Test Only

The statistical-control treatment (test only) was a non-treatment. Students in this treatment did not participate in any motion graphing activities until after the testing was completed. They completed their posttests on the second day of the experiment, when other students were participating in their experimental treatment activities. On the following day, they were encouraged to participate in the MBL activities, while all other students were completing their posttests. Apart from practical and educational considerations, this arrangement helped these students feel equal participants in the research, and motivated them to perform well in their posttests. This treatment was included to control for possible improvements between pretest and posttest simply due to recency and primacy effects.

Subjects

Students participating in the experiment were mostly seniors and their average age was 17.7 years. They came from entire physics classes (with 7-17 students in each class) from seven rural schools in north Florida--a total of 82 students. Rural schools were selected because of their small class size compared with urban schools from several counties. This allowed me to work with entire classes, with minimum disruption to their normal class activities. There are obviously difficulties with statistical interpretation of results from several schools--restriction of random assignment of students to treatment groups, potential lack of equivalence of students from different schools, and threats to external validity. On the other hand, having subjects drawn from a number of schools increases the generalizability of the results.

All students had already been taught kinematics, including relevant conventional laboratory activities, earlier in the academic year, and therefore they should have been familiar with the concepts included in the experimental activities. For this research, I assumed either (a) that students within each treatment had had the same or equivalent exposure to the kinematics content, (b) that any existing differences between students in each treatment were adequately described by the covariates, (c) that knowledge of kinematics concepts was not significantly influence performance on the dependent measure, or (d) that differences in knowledge were randomly distributed across treatments. This assumption is important because random assignment of

students to treatments was restricted. Students could be assigned randomly only to treatment groups within classes, and then in only a restricted fashion when there was an insufficient number.

Students who were absent on either the pretest day or for the treatment activity were automatically assigned to the test-only treatment. This gives some concern, because of the possibility that absentees may not be equivalent to students who attended each day. Absences may result from either lack of motivation (possibly lower ability students) or, on the other hand, because of conflict with other legitimate school responsibilities (higher ability students). Examination of descriptive statistics did not reveal any problems with students who were absent on only one day, so their results were included with the experimental data.

Independent variables

Data provided by students. Students provided information about their age, sex, and number of science and mathematics courses completed. Teachers indicated inconsistencies in some of the reports of science and mathematics courses, and suggested that students may have differed in which courses they included in those numbers (e.g., computer science, agriculture). These values may not be reliable as covariates. Students also provided scores on Scholastic Aptitude Test (SAT) and/or American College Test (ACT).

Verbal test. Because some students had taken neither of these standardized tests, verbal ability was assessed by the French V-1 test (French, Ekstrom, & Price, 1963). This is a timed 8-minute vocabulary matching test. It is widely recognized that verbal ability is a very good predictor of general ability. Sternberg (1986, p. 38) suggested that this is because such tests indirectly measure the ability to acquire meanings of new words from natural contexts and thus reflect ability to acquire new information.

Development. Abstract reasoning and development was measured using a subset of items from the Inventory of Piagetian Developmental Tasks (IPDT, Furth, 1980) as recommended by Patterson and Milakofsky (1980) for use with senior high school students. The administered test contained 8 of the 18 original subtests, each of which has four items. Graphing skills appear to be highly correlated with measures of abstract reasoning, development, and logical thinking (McKenzie & Padilla, 1984; Padilla & McKenzie, 1983).

Performance Measures

Pretest-Posttest

Pretest and posttest were content-specific, dealing with distance-time and velocity-time graphs. No test-retest information is available to determine the reliability of these tests. Every item involved a

physical movement event described in two ways: verbal representation and graphic representation. Students were required to translate from one representation to the other. This translation involved both their understanding of the content, and their ability to interpret graphs as a system of representation. It required performance at the level of "Translation" in Bloom's Taxonomy (Bloom & Krathwohl, 1977). Both pretest and posttest were multiple-choice, objective measures of performance requiring students to understand the underlying properties of concepts and graphs, not just recall of specific events encountered during their activity.

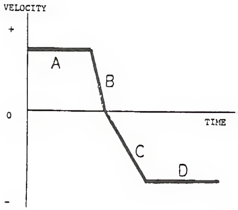
In the posttest, questions were divided clearly on both question booklet and answer sheet into two sections, distance and velocity questions. Each question was on a separate page of the test booklet. Sample questions are shown in Figure 3-1, and the entire posttest is included in Appendix D.

Most of the questions in the posttest involved events of constant velocity, with several questions for each category of constant-velocity events. In recognition of the different ways that students may store their understanding of graphs and concepts of motion, and their different ways of accessing this information, considerable attention was given to constructing posttest questions and alternative answers. Four general categories of questions were formulated, differing on two cognitive dimensions. In addition, three categories of alternative answers were incorporated into each question.

15

Which section of this VELOCITY graph represents an object

moving steadily towards a detector?



20

This VELOCITY graph represents an object

A. with constant positive velocity
B. with constant negative velocity
C. at rest
D. accelerating (increasing velocity)

FIGURE 3-1. Examples of items on posttest showing questions that differ in direction of translation and in conceptual complexity for events of constant velocity.

Question format. The basic task was to translate from one representation to the other. This translation can occur in two directions: from the verbal representation to the graph, and from the graphic representation to the verbal description. These directions of translation were implemented via questions phrased generically as either (a) "Given a verbal description, which graph (or section of a graph) represents it?" or (b) "Given a graph, which verbal description does it represent?"

Verbal description. There are research claims that students may store concepts learned in physics separately from the same concepts learned from the "real" world, and that this may be disguised by competent use of scientific jargon (Clement, 1982; Leboutet-Barrell, 1976; McCloskey et al., 1980). Therefore, it is important to include questions both with and without scientific and mathematical terminology. Parallel questions describe the same event as either (a) "constant positive velocity" or (b) "moving steadily away from the detector." These questions differ in several respects. First, the non-scientific description is considerably more ambiguous. After all, that is the reason for using precise language. For instance, "steady" velocity is less precise than "constant" velocity. It is also not necessary that the detector be the reference point, so walking towards the detector may or may not be positive, depending on the reference system.

Second, the non-scientific description has greater cognitive complexity, requiring students to make two links--one between the scientific concept and the graph, and the other between the colloquial

and scientific descriptions of the event. In addition to the fundamental translation task, such questions require the students to interpret from the imprecise description to the scientific/mathematical equivalent description. Thus, students need to establish that walking steadily implied constant velocity, and that walking away from a reference point (increasing distance) implies positive velocity, before they can address the fundamental task. By itself, this did not seem to create any difficulty. During treatment activities, students unhesitatingly translated instructions to "walk steadily" into actions of constant velocity. However, when such descriptions are coupled with a comprehension task, they may increase the cognitive burden. Scientific terminology or representation is effective only to the extent that students have overlapped real-world operations with scientific terminology.

Alternative answers. Answers were constructed to include, in addition to the correct answer, alternatives that represent common conceptual misunderstandings. One category of answers represented an error simply in the direction of movement (i.e., towards or away from the detector). This indicates a lack of understanding of the graphing conventions and is therefore termed a graph error. A second category of answers indicates that students confused the concepts of distance and velocity, selecting the appropriate distance graph when asked for a velocity graph, and vice versa. For instance, when asked for the velocity graph for an event of constant velocity, students commonly select the graph with positive slope instead of one with zero slope.

This error, referred to in this study as concept error, has been described elsewhere Clement et al., 1986; Mokros & Tinker, in press) as "slope/height" error.

The format of items differed from pretest to posttest in ways that are important in interpreting the results. On the pretest, all the questions involving distance-time graphs were on a single page (both questions and alternative graphs for answers) and all the velocity-time graphs and questions were on a second page (Appendix C). With this pretest format. Students could compare questions and the answers provided. The direction of translation used by the students may not have been the same as the direction of translation intended by the task. In effect, students could make this a matching task, rather than a multiple-choice test.

Graph Test

A computer-administered test of general ability to interpret graphs (Graph Test) was adapted by David Kibbey from an instructional computer program "Relating Graphs to Events" included in "Interpreting Graphs" (Dugdale & Kibbey, 1983). The software uses very few words, and readibility is high. All instructional feedback information was removed from the program. The adapted software recorded the answer selected by each student for each item. In case students' performance on this test exhibited ceiling effects, their decision times (to the nearest 0.1 second) were also recorded. There was no ceiling effect, so the data were not analyzed for that factor. Students were told that their

response time was being recorded, and they were instructed to focus on getting the correct answer, but once they had selected the answer they should enter it promptly.

The 20 items of the graph test were separated into two subtests of 10 items each. One subtest involved speed-time graphs only (speed-graph subtest). The other involved various properties, such as depth, height, area, total income, etc. (miscellaneous-graph subtest). All but one of the graphs contained time as the variable on the absissca. The miscellaneous-graph subtest was used as both covariate (independent variable) as a measure of general graphing ability, and as pretest and posttest (dependent variable) to test whether any improvements in ability to interpret distance and velocity graphs would generalize to other kinds of graphs. The speed-graph subtest was not used as a covariate because of problems of tautology created by using a covariate that was not independent of the distance and velocity graph pretest.

Graph Construction

In the posttreatment testing session, in an attempt to expose students' understanding of the conventions and syntax of Cartesian graphs, students were asked to construct a graph, and in doing so to pay attention to specific features. These attributes, identified by (Padilla et al., 1985) as important components of graph construction skills, were scaling axes, assigning variables to axes, plotting points, and using a best-fit line. Labelling axes was added to these as an additional component. Students were given raw data and asked to draw a

graph to represent the relationship between the height from which a ball was dropped and the height to which it bounced. Scoring these graphs was necessarily subjective, and there is by no means a consensus on the criteria used. These issues will be discussed in more detail when the results are presented in Chapter V.

Attitudinal Measures

Students completed questionnaires both before and after the treatment. Both questionnaires consisted of several Likert-scale items, where positively worded items were alternated with negatively worded ones as far as was reasonable. These questionnaires had not been used before the experiment, so no data are available on the validity and reliability of the items, and results from these questionnaires should be treated with appropriate caution. Results from them are used primarily for explanation and interpretation of the posttest results. The main focus of the questions was perceptions of utility, difficulty, and intrinsic interest. The pretreatment questionnaire was aimed at assessing students' attitudes to graphs (Appendix E), and the posttreatment questionnaire was devoted to determining their attitudes to the treatment activity (Appendix F).

Behavioral Observations

Differences in posttest performance among students in the various treatment groups may be generated as a result of either cognitive or

affective behavioral mediation. A group laboratory activity such as this, which emphasizes discussion and explanation among group members, provides an opportunity to qualitatively examine both verbal and nonverbal behavioral indicators of both types of processes. These observations are an important ingredient in providing adequate explanations of experimental results. These data are, however, predominantly subjective, and should be treated with appropriate caution.

Rogers (1984); Sherman, Webb, and Andrews (1984); Smith (1983); and Wilson (1977) have succinctly reviewed some of the methodological issues that must be addressed for behavioral observation to be a rigorous research techniques. Qualitative research requires that the observer record and report both initial assumptions and subjective reactions to the behavior (i.e., both preconception and postconception). Explanations and descriptions should be obtained through informed selection of data. Observers should try to achieve an adequate formulation through triangulation, convergence (Goetz & LeCompte, 1984), stability, comparability, and plausibility of information.

Preconceptions about student performance and behavior are described within the literature review and in the observation instrument described below. Observations of students' behavior during treatments were recorded in extensive field notes. Several indicators of cognitive and affective mediation were available. Conversation was particularly useful in revealing the level of students' conceptual understanding, the state of their certainty, the intentions of their actions, the strength

of their engagement, etc. It was also useful in demonstrating what kind of procedural and conceptual difficulties were encountered.

All the pilot trials were video recorded, and all treatments were audio recorded. From the pilot trials, an instrument was constructed for rating a limited number of categories of behavior, which could be scored by a sign system at frequent intervals. Audible conversation during treatment activities was scored according to the instrument. Unfortunately, classroom conditions (poor acoustics and a high level of background noise) made it too difficult to derive transcripts from many of the audiotapes. Nevertheless, it was generally possible to detect patterns of behavior.

The categories in the instrument were

(a) orientation--graph (attention to axes, labels etc);
(b) orientation--equipment (how it works, boundaries);
(c) orientation--task (request for information);

(d) procedural--organization (who does what);
(e) procedural--performance;

(f) cognitive--discussion, observation;
(g) cognitive--explanation (reason provided)/prediction;
(h) cognitive--evaluation of results;
(i) cognitive--strategy (planning maneuvers, hypothesizing);

(j) application--reference to school knowledge;
(k) application--reference to out-of-school knowledge;

(l) affective--opt in ("Let me do it")/positive "This is fun";
(m) affective--opt out ("You do it")/negative "This is boring";
(n) affective--aesthetic, kinesthetic;
(o) affective--doodling (constructing graph as picture etc);
(p) affective--off task.

Experimentally-Derived Artifacts

Throughout the MBL experience, students followed structured worksheets where they were regularly asked to "Describe (in words or sketch) what happens to the graph when you . . . ." Their choice of using graph or words can be compared with their response on the pretreatment questionnaire to the item asking whether they voluntarily use graphs. Their responses to supplementary questions on the worksheet can provide explicit information about their understanding of key concepts during the activity.

Statistical Analysis

Unit of Analysis

Except for students in the test-only treatment, all students participated within groups for the treatment activity. The basic research design was to have one group for each treatment (as far as class size would permit) within each school. This meant that, within each treatment, there were two levels of organizational units available for analysis of results--the individual student and the group (or class, or school). The group contexts may have influenced the behavior, thoughts, and attitudes of the group members.

In the last decade there has been considerable interest in the appropriate level to use for statistical analysis of empirical experimental results, and how that analysis should be performed. This

debate has been reviewed by Burstein (1980), Cronbach (1976), and Raudenbush and Bryk (1986). Some claim that, because pupils react differently in the group than by themselves, it is appropriate to use the group as the unit of analysis for the evaluation of programs where instruction is received simultaneously by all students in the class. Counterarguments claim that pupils react as individuals and thus the effects on them should be the focus of the evaluation.

Phenomena of importance occur at all levels of the system, and they need to be described and investigated. Analyses involving both individual-level and group-level effects are important, and should be based on theories in which the source and form of group effects are stated specifically. A single level analysis is seldom appropriate for multilevel educational data. In conducting a contextual analysis it is useful to know whether aspects of group membership contribute to an explanation of individual outcome behavior after the effects of individual characteristics have been considered.

There are many ways of measuring and analyzing group data, but typically, group effects are analyzed after aggregating the observations of the members of the group on some variable of interest. The same variables can be used for analysis at different levels, but they can have different meanings at each level. Group-level analyses minimize problems with unequal group size and missing data, assuming that attrition and assignment to groups are random. They also reduce concerns about measurement errors, but, on the other hand, the severe reduction in degrees of freedom often makes it impossible to do any more than broad-scale analyses at the group level.

Contextual analysis, which is the study of the effects of group-level (macrolevel) variables on individual-level (microlevel) outcomes, suggests two main sociological or social psychological explanations of context effects. The first are normative effects, where group values or behavior establish a normative climate which influences individual motivation to learn, or where group average effects mediate achievement through informal interpersonal processes. The second are comparison effects, where the group value on a characteristic functions as a comparison point against which an individual can evaluate himself (like a big frog in a small pond or a little frog in a big pond).

In this research, performance measures were statistically analysed with both the individual and the group as the unit of analysis. For the group analysis, data were aggregated from the data for students in each group to provide a group mean for each variable. Group means then became the basic units of information for that variable. The basic statistical procedure was a generalized linear model of a factorial analysis of covariance, using the Statistical Analysis Systems (SAS) library.

Analysis of Performance Measures

Initial descriptive statistics involved determining means for the group of interest (treatment, class, sex), and determining correlations and regressions between covariates. A probability level of .05 was used for statistical analyses of performance measures. Comparisons of interest were between the following groups: standard MBL and delayed

MBL, standard MBL and pencil and paper, delayed MBL and pencil and paper, and pencil and paper and test only.

Information on covariates was analyzed to determine whether there were significant differences among students (or groups) assigned to each treatment. Stepwise regression determined which covariates were significant predictors for each outcome measure. In no case was there a significant interaction between the covariates and the grouping factors (treatment and class). A factorial analysis of covariance was employed to determine treatment effects because it is effective in reducing the error term. Gender differences were of interest in this study, and were included as an independent variable in separate covariance analyses.

The residuals from this model were examined, and in no case did they indicate any serious breach of the assumptions of conditional normality, linearity, and homoscedasticity. The assumption of homogeneity of regression slopes was checked during the test for interaction between pretests and treatment.

Analysis of Attitude Measures

After performing basic descriptive statistics, a correlation matrix was employed to determine whether items on each questionnaire were independent. Items that were similar in both construct and response were aggregated for further analysis of variance to determine treatment or gender differences. In this analysis, a probability level of 0.10 was adopted for two reasons: the acknowledged tendency to respond with

the central value, and the wording of the extreme values on the Likert scale was sufficiently extreme to discourage students from using them.

Summary

The research concentrated on real-time graphing as the most salient feature of the MBL tool. The impact of real-time graphing of the standard-MBL treatment was isolated by delaying the graph display for 20-30 seconds (by altering the software) but leaving the activity the otherwise identical. These two MBL experimental treatments were compared with a pencil-and-paper graphing activity and no activity. The basic research design involved one group of each of four treatments within each of seven small physics classes. Because of differences in class sizes and student absences during the experiment, this design was not completely achieved, nor was it balanced (Table 3-2).

This research incorporates the usual battery of independent variables (Table 3-1) used to explain variance in many kinds of learning and performance (i.e., sex, age, general ability, reasoning and development). However, the literature does not provide a valid, reliable instrument for assessing graphing skills and revealing the underlying concepts that students have about graphs and graphing (i.e., their graph schema). One reason for this is the difficulty of separating the conceptual understanding of graphs, as a system of representation, from that of the concepts being represented. In addition to a content-specific posttest interpreting graphs of motion, students completed tests of ability to interpret graphs of a wide range

of physical phenomena and a test of ability to construct a graph from experimental data provided for them. Results from these tests provides guidance for explaining experimental results and for designing future research. Students also completed questionnaires to assess their attitudes to graphs (pretreatment) and to the treatment activity (posttreatment). Results were analysed using analysis of covariance.

The research extended over three days: (a) pretreatment, (b) treatment, and (c) posttreatment. Including students who participated in the pilot study and are included for pretest data only, there were 93 high school physics students in the study. The single class period was intended as an experiential treatment, not an instructional one. It is considered to be too short for significant remediation of stable misconceptions about distance and velocity.

CHAPTER IV

POSTTEST--RESULTS AND DISCUSSION

Introduction

The basic research design tested for differences in performance among four treatments--standard MBL, delayed MBL, pencil and paper, and test only, administered to small groups of students at each of seven, small, rural schools. The numbers of students in each group were not equal (Table 3-2). The covariates and performance measures were described in the previous chapter (Table 3-1). Where pretest data only were being analyzed, data were included from students who participated in the pilot study under the same testing conditions.

The research results are analyzed in three sections. In this chapter, results are presented from analyses of performance on the kinematics pretest and posttest, first with the individual, then with the treatment group, as the unit of analysis. Responses are examined for indications of the kinds of problems students had with either the concepts or the graphs, and also for which of these problems were addressed by the treatments. The following chapters will present experimental results as they provide information about graphing skills (attitudes and problems) (Chapter V), and attitudinal and behavioral

indicators of information processing and motivation during the treatment activity (Chapter VI).

Descriptive Statistics of the Total Population

Descriptive statistics were obtained for all students participating in the research ($\underline{N}$=93), including the 11 senior students in the pilot study and students who were absent for the posttest, but not including the tenth grade students who contributed to the pilot study. Table 4-1 (using abbreviations for covariates and test scores as described in Table 3-1) provides a summary description of these measures on covariates and pretests. Correlations among covariates and pretests are shown in Table 4-2. Throughout this section a critical probability level of .05 was taken to indicate statistical significance.

Scores from the three measures of general ability--SAT, ACT, and Verbal--were highly correlated. Of these measures, SAT scores are the most widely known, so they were used as the reference ability measure. Where SAT scores were unavailable (54 students, equally distributed among treatments), they were estimated by regression using both ACT and verbal scores ($\underline{n}$=35, $\underline{r}^2$=.826, $\underline{p}$<.001) or by regression using verbal score alone ($\underline{n}$=19, $\underline{r}^2$=.438, $\underline{p}$<.001). This single score of measured or estimated SAT was used in all statistical analyses.

Not surprisingly, the numbers of science and mathematics courses were positively correlated with the age of the student. Generally, students add math and science courses as they go through high school at the rate of approximately one course of each subject each year. On the

TABLE 4-1. Descriptive statistics (mean, and standard deviation in parentheses) of the student participants in each treatment group.

Covariate	Standard MBL (n=18)	Delayed MBL (n=19)	Pencil-Paper (n=18)	Test Only (n=20)
Age	210.44 (4.67)	211.26 (6.88)	212.33 (4.97)	213.50 (8.42)
Science	4.11 (0.90)	3.92 (1.17)	3.97 (0.61)	3.68 (0.80)
Math	3.69 (0.67)	3.95 (0.86)	4.36 (1.00)	3.38 (0.81)
SAT	1049.2 (134.1)	913.2 (154.9)	1014.3 (141.1)	986.2 (118.8)
Development	24.33 (3.03)	23.21 (4.38)	24.89 (3.45)	24.25 (3.60)
Graph Test	15.67 (2.72)	14.58 (2.91)	14.56 (2.73)	14.15 (2.64)
Graph-Misc	8.00 (1.14)	7.58 (1.68)	7.22 (1.52)	7.35 (1.69)
Graph-Speed	7.67 (1.91)	7.00 (1.56)	7.33 (1.88)	6.80 (1.54)

TABLE 4-2. Pearson correlation coefficients among covariates and pretest measures (N=93), and probability of r>0 in parentheses.

	Age	Sci	Math	SAT	DEV	Graph	PRE
Age		.233 (.025)	.245 (.018)		.204 (.050)		
Sci				.346 (.001)		.346 (.001)	
Math		.199 (.057)					
SAT	.008 (.937)		.075 (.484)		.381 (.001)	.529 (.001)	.333 (.001)
DEV		.180 (.085)	.173 (.099)			.470 (.001)	
Graph	.192 .065		.172 (.101)				.454 (.001)
PRE	.086 (.415)	.112 (.286)	.078 (.460)		.108 (.301)		

Significant coefficients ($p<.05$) are shown above the diagonal, and coefficients that do not reach this critical level of significance are shown below the diagonal.

average, students in this study had one more year of science and one more year of math than was required for graduation under the 1986 state standards for Florida.

Students who take physics courses in high school are generally among the brightest of the students. This is reflected in a reasonably high mean SAT score. Considering that the development and reasoning test consisted of a subset of the most difficult items in the IPDT (described in the methods chapter), the development scores probably

suggest that most of the students were at least transitional between concrete and formal operational levels according to Piaget's stages of development. This exclusivity of students in physics courses means that the range of ability of students in this research is restricted.

School Differences

Each school differed in the degree of exclusivity of the physics students. The differences in natural selection of physics students are a consequence of community characteristics, the quality and motivation of the students, the quality and availability of teachers, and the type of counselling. In this study of students from nine schools, physics students as a proportion of graduating seniors ranged from 4% to 34%. Not surprisingly, therefore, there were significant differences among students from the various schools in age, number of science and mathematics courses completed, reasoning and development (IPDT), and in general ability to interpret graphs (Graph Test).

Equivalence of Treatment Groups

At each school, students were assigned randomly to different treatment groups. With one exception, scores from students within each treatment were not significantly different on the covariates or pretest measures. The sole exception was SAT (Table 4-1), where scores of students in the delayed-MBL treatment were significantly lower than those of students in both the standard-MBL and the pencil-and-paper

treatments. There were disproportionately more females in the delayed-MBL group (12 females, 7 males). Because females generally had lower SAT scores than males (discussed later in this chapter), the data were analyzed to determine whether the low SAT scores could be accounted for by the high number of females in this group. After taking account of gender, differences in SAT scores remained (Table 4-3), which were significant between delayed-MBL treatment and both standard-MBL (p=.006) and the pencil-and-paper (p=.049) treatments.

The effect of this pretreatment difference among treatments in SAT scores would be to make posttest means adjusted by this covariate higher than they ought to be for the group higher in SAT scores and lower than they ought for the group low on SAT. The bias is in the direction of exaggerating differences on posttest scores adjusted by this covariate between groups high on both SAT and posttest and groups low on both measures. Thus it poses some concern of artificially creating significant differences between the standard-MBL (high scores on both SAT and posttest, see results later in this chapter) and the delayed-MBL

TABLE 4-3. Regression analysis to determine whether differences among treatments on SAT scores could be explained by differences in sex ratio.

VARIABLE	df	TypeIII SS	F	p	r^2
Dependent variable: SAT					
Model	4	247,620	3.37	.014	.163
Sex	1	61,422	3.34	.072	
Treatment	3	162,773	2.95	.039	
Error	69	1,515,981			

treatments (lower scores on both measures). This nonequivalence could also diminish differences between groups high on only one of these measures (i.e., between delayed-MBL and pencil-and-paper treatments).

Pretest Performance

Scores on the kinematics pretest were significantly correlated with ability measures of SAT (r=.333, p=.001), and with the graph test (r=.454, p=.001). They were not, however, significantly correlated with development (r=.108, p=.301). There were too few items on the pretest, and hence the variance was too high, for detailed analysis of performance on the pretest. The following equation was generated from stepwise regression as the best model, explaining 17% of the variance of scores on the pretest.

Pretest = 3.005 + 0.421 Graph-Speed [r^2=.177, p<.001]

Posttest Performance

Individual as the Unit of Analysis

The kinematics posttest was combined from two subtests--distance subtest (Post-D) with 8 distance-time graph items, and velocity subtest (Post-V) with 16 velocity-time graph items. Scores on the subtests were not significantly correlated with each other for any of the treatment groups (standard MBL r=.165, p=.51; delayed MBL r=-.206, p=.40; pencil and paper r=.071, p=.78; test only r=.083, p=.73). Results throughout

this dissertation reinforce the image of a qualitative difference between the concepts of distance and velocity. Performance on each of these subtests (Figure 4-1) can be compared with the pretest scores in Table 4-4. Because of differences in format between pretest and posttest (see methods chapter), scores on the two tests cannot be compared directly to obtain a measure of change in performance. These issues will be discussed in more detail when the test items are analyzed later in this chapter.

Stepwise regression indicated the following combination of covariates, each contributing significantly to explaining the variance in each dependent variable.

Posttest = 0.982 + 0.0098 SAT + 0.794 Pretest [r^2=.380, p<.001]

[SAT partial r^2=.242, p<.001; Pretest partial r^2=.139, p<.001]

Post-D = 0.095 + 0.0037 SAT + 0.322 Pretest [r^2=.209, p<.001]

[Pretest partial r^2=.136, p<.002; SAT partial r^2=.073, p=.013]

Post-V = 1.067 + 0.0051 SAT + 0.469 Graph-Misc [r^2=.186, p<.001]

[SAT partial r^2=.139, p=.001; Graph-Misc partial r^2=.047, p=.047]

Not surprisingly, SAT was the most significant covariate for predicting posttest performance on both subtests. The other significant covariate was either the pretest score (for the distance subtest) or the speed-graph score (for velocity subtest) which was significantly correlated with pretest performance (r=.425, p<.001).

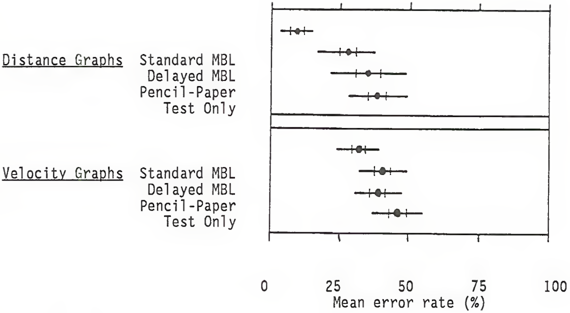


FIGURE 4-1. Mean posttest scores for the distance and velocity subtests. Mean, 50%, and 95% confidence limits are shown.

TABLE 4-4. Mean scores for each subtest of pretest and posttest.

Treatment	n	Distance Mean	Distance s.d.	Velocity Mean	Velocity s.d.	Total Mean	Total s.d.
		Pretest					
		(4 items)		(5 items)		(9 items)	
Standard MBL	18	3.22	1.06	3.17	1.25	6.39	1.79
Delayed MBL	19	3.32	1.06	2.89	1.15	6.21	1.58
Pencil-Paper	18	3.44	0.92	2.50	1.25	5.94	1.62
Test Only	20	3.20	1.20	2.60	1.57	5.91	2.02
		Posttest					
		(8 items)		(16 items)		(24 items)	
Standard MBL	18	7.28	0.89	10.94	2.51	18.22	2.80
Delayed MBL	19	5.79	1.75	9.58	2.87	15.37	3.04
Pencil-Paper	18	5.17	2.15	9.83	2.87	15.00	3.71
Test Only	20	4.95	1.79	8.60	3.05	13.70	3.57

After controlling for significant covariates (determined by the stepwise regressions described above), factorial analysis of covariance indicated significant differences ($p<.001$) among treatments in the posttest (Table 4-5). Of the comparisons of interest, the ones that were significant ($p=.05$) were between the standard-MBL and both delayed-MBL and pencil-and-paper treatments.

TABLE 4-5. Analysis of covariance for posttest scores with the individual as the unit of analysis. Covariates were determined by stepwise regression.

VARIABLE	df	TypeIII SS	F	p	r^2
Dependent variable: Posttest					
Model	5	491.78	14.09	<.001	.533
SAT	1	99.93	14.32	<.001	
Pretest	1	93.11	13.34	<.001	
Treatment	3	138.04	6.59	<.001	
Error	68	474.45			
Dependent variable: Distance Posttest					
Model	5	100.93	8.23	<.001	.377
Pretest	1	12.86	5.25	.025	
SAT	1	13.51	5.51	.022	
Treatment	3	47.57	6.47	<.001	
Error	68	166.69			
Dependent variable: Velocity Posttest					
Model	5	152.24	4.42	.002	.245
SAT	1	25.27	3.67	.060	
Graph-Misc	1	21.03	3.06	.085	
Treatment	3	37.06	1.80	.156	
Error	68	467.98			

Thus the real-time graphing feature was effective in improving posttest performance, and other features of the MBL were not, when compared with the pencil-and-paper reference activity. Mean scores of the different treatments indicate that real-time graphing (i.e., the difference between standard-MBL and delayed-MBL scores) accounted for nearly all (90%) of the improvement relative to the control (i.e., the difference between standard-MBL and pencil-and-paper scores). However, the comparison between standard- and delayed-MBL treatments may be statistically exaggerated by the bias introduced by nonequivalence of students' SAT scores within each treatment (see earlier discussion).

At no time throughout this research was there evidence of any difference between scores from the pencil-and-paper students and those from the test-only students. It appears that, if there was any improvement in performance of pencil-and-paper students, it was no greater than that resulting from recency and primacy effects.

The treatment effect can be seen more clearly for the two subtest scores (Figure 4-1). The treatment effect was significant for the distance subtest (p=.002), but not for the velocity subtest. For the distance subtest, the standard-MBL students again scored significantly higher (p=.05) than both the delayed-MBL and pencil-and-paper students. Regression data for these analyses are provided in Table 4-5.

Treatment Group as the Unit of Analysis

A similar series of analyses was performed on data aggregated for students who worked together as a group for the treatment activities.

This resulted in a maximum sample size of seven for each treatment (one each at each school). The number of students in each group was not equal (Table 3-2). Because of the very small sample size, and the unbalanced number of students in each group, there is more concern with a Type II error than with a Type I error, and a probability level of .10 was used to determine statistical significance for these analyses. The sample number was also too low for any tests of interaction.

After controlling for significant covariates (miscellaneous-graph subtest) determined by stepwise regression, factorial analysis of covariance indicated significant differences among treatments in the distance subtest ($p<.001$), but not in either the velocity subtest or the total posttest (Table 4-6). These results are generally consistent with the results from analyzing the data from individual students. Again, pairwise comparisons of treatments were significant between the standard-MBL and all other treatments.

Item Analysis and Discussion of Posttest Performance

Even though the analysis of covariance demonstrates a significant treatment effect, a cursory comparison of pretest and posttest results (Table 4-4) suggests a decrease of about 20% in performance of the pencil-and-paper and test-only groups, rather than an increase in performance of the students in the two MBL groups. However, the pretest and posttest were not the same (see details in the methods chapter) and cannot be compared directly in this way. Although the same concepts were being tested, substantive differences in question format (which

TABLE 4-6. Analysis of covariance for posttest scores with the group as the unit of analysis. Covariates were determined by stepwise regression.

Variable	*df*	TypeIII *SS*	*F*	*p*	r^2
Dependent variable: Posttest					
Model	4	103.13	8.77	<.001	.626
Graph-Misc	1	30.95	10.53	.004	
Treatment	3	19.69	2.23	.114	
Error	21	61.73			
Dependent variable: Distance Posttest					
Model	4	22.83	5.71	.004	.507
Graph-Misc	1	3.08	2.91	.102	
Treatment	3	10.01	3.15	.046	
Error	21	22.22			
Dependent variable: Velocity Posttest					
Model	4	34.48	2.80	.052	.348
Graph-Misc	1	14.75	4.79	.040	
Treatment	3	3.82	0.41	.745	
Error	21	64.61			

were discussed in the previous chapter) may have influenced the difficulty of the items.

Many of the items in the pretest and posttest involved movement events of constant velocity--two distance graphs and two velocity graphs in the pretest, and six distance graphs and seven velocity graphs in the posttest. In the posttest, these items were designed to differ in question format (direction of translation between verbal and graphic descriptions of each event), verbal description (scientific or nonscientific terminology), and alternative answers (representing

"graph" or "concept" errors). These differences were detailed in the methods chapter. Performance on constant-velocity items can be examined for information about what conceptual problems students had and what they learned.

In the posttest, where students were asked equivalent questions in different ways, they gave different answers. Whitaker (1983) found that, in addition to giving different answers, some of his students gave different reasons for selecting a given answer. Stability of students replies, question format, and instructional context may well be factors here. Because of the differing structure and format of questions dealing with equivalent concepts, particularly in the posttest, it is possible to analyze the items for the influence of question format and item complexity on error rates. Representative items were shown in Fig. 3-1. Unfortunately, the small number of items involved and the unbalanced nature of the items in each category do not allow rigorous statistical analysis. Descriptive statistics only (group means) are provided, and the results are examined for consistent trends.

Item Analysis

Direction of translation. As detailed in the methods chapter, question formats differed in two conceptual dimensions. The first of these is the direction of translation (between verbal and graphic symbolic descriptions of an event). In most items students were given a verbal description of a motion and asked to select the appropriate section of a graph to represent it (Question 15, Figure 3-1). This

required them to translate from verbal to graphic representation of the event (verbal => graph). Other items provided a graph and asked students to select the appropriate verbal description of the event, requiring them to translate from graphic to verbal representations of the event (graph => verbal) (Question 21, Figure 3-1).

Cognitive complexity. Additionally, events were described at two levels of cognitive complexity. Items that described the movement event as "constant positive (or negative) velocity" tested the students' ability to translate directly between two systems of symbolic representation (verbal and graphic) of a specific concept. Such items are referred to here as translation (T) items. In other items, the same event was described as movement "steadily away from (or towards) a detector." Before they could make the appropriate translation, students had to interpret the verbal description in terms of the fundamental concepts, or translate from colloquial to mathematical-scientific verbal descriptions. They had to realize that steady movement implies constant velocity, and that movement away from a reference point results in increased distance. These items are referred to as interpretation + translation (I+T) items.

Posttest results for velocity-time graphs are broken down into these categories of items in Figure 4-2 and Table 4-7. Within each category, there is a fairly consistent treatment effect, with the percentage errors decreasing from the pencil-and-paper, to the delayed-MBL, to the standard-MBL treatments. Apart from the first category,

translating from graph to verbal description for which there was only one item, these differences are not particularly striking.

Within each treatment group, the mean error rate was greater for items requiring translation from verbal descriptions to graphs than for than items requiring translation in the other direction. It was also greater for I+T items than for items requiring translation alone. That is, the interpretation task placed an additional cognitive burden on the students. Similar patterns were apparent in error rates for distance graphs of constant velocity events, and for velocity graphs dealing with an event where the object reversed direction (turned around).

TABLE 4-7. Mean error rate (MER, %) for questions of velocity graphs differing in the direction of translation between verbal and graphic descriptions (V=>G or G=>V), and in the complexity of the task--translation (T) or interpretation plus translation (I+T).

			------------ Treatments ------------			
Direction of Translation	Task	# items	Standard MBL	Delayed MBL	Pencil Paper	Test Only
G => V	T	1	6	21	33	25
G => V	I+T	1	33	47	50	60
V => G	T	2	44	58	53	62
V => G	I+T	2	67	74	84	88
Difference in Mean Error Rate by Reversing Direction of Translation (i.e., MER (V=>G) - MER (G=>V))						
	T		38	37	20	37
	I+T		34	27	34	28
Difference in Mean Error Rate by Changing the Task Complexity (i.e., MER (I+T) - MER (T))						
G => V			27	26	17	35
V => G			23	16	31	26

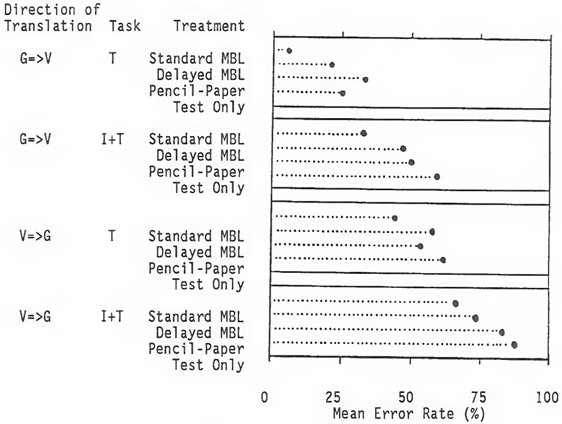


FIGURE 4-2. Mean error rate (%) for questions of velocity graphs differing in the direction of translation between verbal and graphic descriptions (V=>G or G=>V), and complexity of task--translation (T) or interpretation plus translation (I+T).

These results suggest that students had errors associated with (a) linking the graph and verbal representations (T items) and (b) linking verbal descriptions in scientific and colloquial terminology (difference between I+T and T items). There is no indication of any treatment effect with this latter category of linking errors (i.e., (I+T)-(T)). The standard-MBL experience does not seem to have improved the students' linkage of real-world events (or at least descriptions of events) with the relevant scientific terminology.

We can think of the translation between verbal and graphic symbol systems as being analogous to translation between two verbal languages. As practical experience tells us that it is easier to translate from a less familiar language to a more familiar one than the reverse, so it is reasonable that students should have less difficulty translating from graphs (a less familiar system of representation) to words (a more familiar representation), than from words to graphs. It would be interesting to see whether performance for translating in each direction would approach parity as students improve their graphing competence.

In related research with the motion detector with middle-school students, Mokros (personal communication, May, 1986) found that students had a higher error rate with a multiple choice item than when they were asked to sketch the appropriate graph. She suggested that, when students are provided with an incorrect graph representative of the most common misconceptions, they find it so visually compelling that they select it without really thinking about it. Different types of mental processes that are called into play by the different test contexts suggest that what may be tested in the multiple-choice context is the availability of knowledge rather than the knowledge per se.

Questions requiring translation in different directions were not equivalent. The alternative answers were presented as sections of a single graph rather than as separate graphs. It is possible that apparently subtle differences, such as this, could have affected students' ability to answer multiple choice questions. The profound influence of question format on performance described in this section raises the question of whether the students are skilled in using cues

from question format rather than in knowing the concept that is supposedly being tested.

Response Analysis

What did the students in the standard-MBL treatment learn--to follow conventions, to interpret graphs, or to understand the concepts of kinematics? An analysis of students' responses on the constant-velocity items in the posttest can provide some clues. The treatment period was too brief to expect significant progress in remediating reportedly stable misconceptions about distance and velocity. However, behavioral observations (which will be discussed in a later chapter) indicated that they were attending to, and presumably learning, attributes of graphs and concepts of motion.

Most of the items dealing with events of constant velocity included options that indicated two common types of errors. In the first, students erroneously selected the graph for positive velocity when given an event of negative velocity, and vice versa. This confusion between positive and negative velocity, or between directions of the physical movement, probably results from not fully understanding the conventions of graphic representation. Such errors are therefore termed "graph" errors. However, to some extent, they may also represent a failure to distinguish between scalar and vector properties (i.e., between speed and velocity). In the second category of errors, students erroneously select a flat distance graph or a sloping velocity graph for events in constant velocity. This probably results from conceptual confusion on

the concepts of distance and velocity. Such errors are termed "concept" errors. Students' responses are categorized in this way in Figure 4-3 and Table 4-8. There is no way of determining to what extent the selection of alternative answers may have reflected students' nonunderstanding (i.e., all or partly guesswork).

Most of the improvement by standard-MBL students was a reduction in graph errors (Figure 4-3). Graph error rates for distance graphs were significantly lower ($p<.05$) for standard-MBL students than for students from all other treatments. For velocity graphs, the improvement was less marked, with standard-MBL students having fewer graph errors than students in both pencil-and-paper and test-only treatments ($p<.10$), but not having significantly fewer errors than the delayed-MBL students.

Even though the treatment period was brief, it is not surprising that the opportunity for practice afforded by the standard-MBL treatment should have improved students' understanding of the conventions of graphic representation. What is surprising is that the short delay in displaying the graph in the delayed-MBL treatment cancelled the effect.

Students in all activity treatments showed improvement with distance concept errors compared with the test-only students (both MBL treatments $p<.05$, pencil and paper $p<.10$). The concept of velocity provided the greatest source of difficulty on both pretest and posttest, and there was no evidence of improvement with this concept for any of the treatments. Students persisted in selecting a sloping velocity graph instead of a flat graph to represent a constant-velocity movement for an overall mean of 44% of the items.

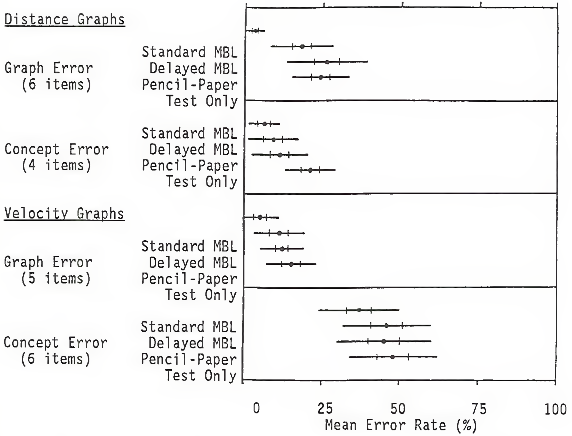


FIGURE 4-3. Mean error rate (%) for different responses (representing graph errors and concept errors) to posttest items involving events of constant velocity. Mean, 50%, and 95% confidence limits are shown.

TABLE 4-8. Mean error rate (%) for responses to items involving events of constant velocity, indicating two common conceptual errors.

Type of Error	# items	Standard MBL	Delayed MBL	Pencil Paper	Test Only
		Distance Graphs			
Graph	6	3	17	27	26
Concept	4	6	12	11	22
		Velocity Graphs			
Graph	5	4	10	12	15
Concept	6	32	40	34	32

The higher proportion of concept errors with velocity graphs than with distance graphs indicates that velocity is conceptually more difficult than is distance. Velocity is an more abstract property, being derived as the rate of change of distance. The lack of improvement with the concept of velocity was anticipated because of the brief treatment period and the evidence in the literature indicating that students' misconceptions and confusions between distance and velocity (or speed) are widespread and stable, being resistant to countersuggestion and conventional instruction.

Resilience of Conceptual Errors

Stability has two key components--resistance (not being easily moved) and resilience (return to the original state after being moved, like elasticity). Most studies in the literature have examined only the resistant properties of misconceptions. What has not been so widely acknowledged is the resilience of these alternative concepts.

In this research, it is possible to look at resilience of the conceptual confusion between distance and velocity by comparing students' responses to supplementary questions on the treatment worksheets with their responses to questions on the same concepts in the posttest. There are three concepts available to compare in this way: (a) for a constant velocity event, the distance graph should have a positive slope; (b) for a constant velocity event, the velocity graph should be flat; and, for MBL groups only, (c) when an object reverses direction, the instantaneous velocity should be zero.

Error rates on the posttest were low for items involving distance graphs and will therefore not be examined further, but errors were high for both of the other concepts. In many cases, students answered these questions correctly on their worksheets, but made errors with the very same concept on the posttest the next day (Figure 4-4).

There are two reasons for caution in interpreting the results in this way. The first is based on the impact of question format on error rates, as discussed above. The second focuses on the social environment of the treatment activity when students responded to adjunct questions in their worksheets. Within each group, most students entered the same responses to each question in the worksheet. On the posttest, there was at least one student from each group who made none or few of these errors. Given these facts, it seems likely that the responses on the worksheets of the entire group may well represent the understanding of this group member whose understanding of the subject matter was the most complete. Transcripts of recorded conversation among students during their treatment activities indicated that there was often considerable discussion on these points during the treatment activity, and students seemed to believe the answers they wrote at the time they wrote them. A single class period was apparently inadequate to consolidate their altered conceptions.

There was another reason for the lack of improvement in velocity graphs. During the MBL treatment, the usual sequence of events was for one student to start the computer to gather data, then another student would start moving. It took this student a few seconds to reach steady speed, without any of the group being really conscious that the mover

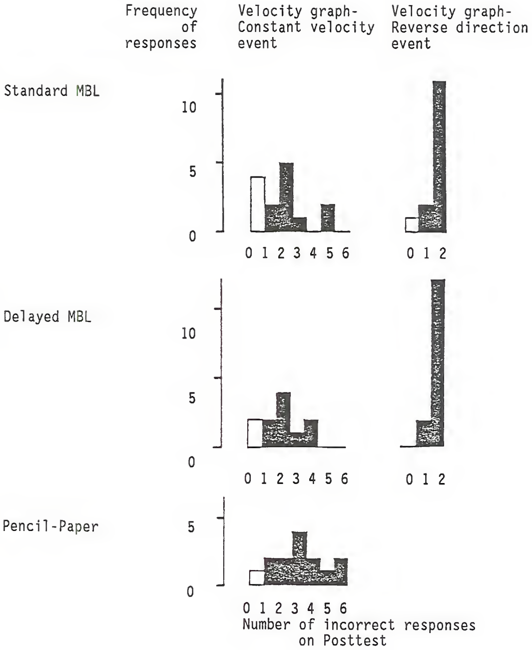


FIGURE 4-4. Resilience of students' misconcepts about velocity--frequency of errors on posttest after correctly answering questions during the treatment activity. White bars indicate no errors on posttest.

was actually accelerating. As the mover accelerated, the velocity graph had a nonzero slope. If they were trying to move fast, they may even have been out of range of the detector before they reached their steady speed. They expected to see a sloping graph, and they did see a sloping graph. By attending to the wrong part of the velocity graph, their misconceptions were consolidated, not refuted. It was unfortunate that very few of the students completed the third phase of the designed activities, where students perform the movements necessary to reproduce a given graph. In pilot studies, this was where students were confronted with a conflict or discrepancy between their expectations and reality. This was where their misconceptions appeared to be overcome.

Gender Differences

Because of the interest in potential differences between sexes in spatial skills, the descriptive data have been separated in Table 4-9. Considering past sex bias in high school physics classes, the female students were gratifyingly well represented, comprising 47% of the research population. Females were generally younger, and had lower ability and lower development scores than did males. Since there were significant intercorrelations (Table 4-2) among age, SAT, and reasoning, regression analyses were performed to determine whether the differences between sexes on SAT and development could be explained by the differences in age (Table 4-10). After taking account of age differences, differences in SAT remained (p=.001) but differences in development were not significant (p=.094).

TABLE 4-9. Sex differences on covariate measures and on pretests for all students (including students not included in posttest analyses). Standard deviations are given in parentheses.

Variable	Females (n=44) Mean	s.d.	Males (n=49) Mean	s.d.	sig.
		Covariates			
Age (in months)	210.36	(5.46)	213.73	(7.04)	**
SAT	948.5	(147.1)	1030.5	(136.0)	**
DEV	23.55	(3.97)	25.16	(3.31)	**
		Pretests			
Graph Test	14.18	(2.87)	15.67	(2.35)	**
Graph-Misc	7.43	(1.37)	7.78	(1.61)	
Graph-Speed	6.75	(1.88)	7.90	(1.28)	**
Pretest	5.80	(1.41)	6.41	(1.88)	*
Pre-D	3.36	(1.01)	3.35	(1.05)	
Pre-V	2.43	(1.15)	3.06	(1.31)	**

Significance levels: * p<.10, ** p<.05.

Sex differences were also apparent on both the kinematics pretest and the graph pretest, with males having higher scores than females. On each of these tests the sex differences were significant only for the subtests involving speed-time items (graph pretest) or velocity items (kinematics pretest). The consistency of these two results suggests a qualitative difference between the concepts being graphed (distance and velocity), probably in conceptual complexity or degree of abstractness.

Females may have had a poorer conceptual understanding of velocity, a concept which is more complex and more abstract than distance. After

TABLE 4-10. Analysis of covariance to determine whether the differences between sexes on SAT and development scores could be attributed to differences in age.

Variable	*df*	TypeI *SS*	*F*	*p*	r^2
Dependent variable: SAT					
Model	2	156831	3.88	.024	.082
Age	1	138	0.01	.934	
Sex	1	156693	7.76	.001	
Error	89	1757360			
Dependent variable: Development					
Model	2	90.10	3.45	.036	.071
Age	1	52.65	4.04	.048	
Sex	1	37.45	2.87	.094	
Error	90	1174.18			

controlling for differences in SAT and development (Table 4-11), the sex difference was not statistically significant for the velocity subtest (p=.08), although it remained significant for the speed subtest (p=.04).

Posttest results were analyzed for sex by treatment interaction (Table 4-12). This interaction was significant only for the distance subtest (p=.044). After separating the results for each treatment and controlling for differences in covariates, there were significant sex differences on the kinematics posttest for students in the standard-MBL treatment (Table 4-13). For the distance items, females scored significantly higher than did males (p=.025), while the reverse was true for velocity items (p=.048). Both males and females scored about 90% on the distance subtest, so ceiling effects may have limited the detection

TABLE 4-11. Analysis of covariance to determine whether sex differences on pretest measures could be attributed to differences in SAT and development.

Variable	df	TypeI SS	F	p	r^2
Dependent variable: Graph-Speed subtest					
Model	3	79.22	12.56	<.001	.305
SAT	1	16.55	7.87	.006	
DEV	1	16.38	7.79	.007	
Sex	1	9.12	4.34	.040	
Error	86	180.78			
Dependent variable: Pretest-Velocity subtest					
Model	3	16.30	3.59	.017	.111
SAT	1	6.98	4.60	.035	
DEV	1	0.15	0.10	.757	
Sex	1	4.68	3.09	.082	
Error	86	130.32			

of improvement for males (mean score of 86% on the distance items in the kinematics pretest). In general, groups with low error rates also had comparatively low variances associated with the mean score. There were no sex differences for students in the pencil-and-paper treatment.

The results can also be examined for differential treatment effects between the sexes. In Table 4-14 the data have been separated by sex, and after controlling for significant covariates, the subsequent factorial analysis of covariance indicates a treatment effect for females for distance subtest only (p=.002). Females in the standard-MBL treatment made fewer errors on the distance subtest, than did females in the pencil-and-paper treatment. The treatment effect for males was not

TABLE 4-12. Analysis of covariance to determine sex by treatment interaction on the posttest.

Variable	*df*	TypeIII *SS*	*F*	*p*	r^2
Dependent variable: Posttest					
Model		509.33	7.92	<.001	.527
SAT	1	86.58	12.12	<.001	
Pretest	1	81.15	11.36	.001	
Treatment	3	132.20	6.17	.001	
Sex	1	4.14	0.58	.449	
Sex x Treat.		13.53	0.63	.597	
Error		457.12			
Dependent variable: Posttest-Distance					
Model		120.90	5.86	<.001	.452
SAT	1	13.91	6.07	.016	
Pretest	1	14.74	6.43	.014	
Treatment	3	44.07	6.41	.001	
Sex	1	0.34	0.15	.703	
Sex x Treat.		19.70	2.86	.044	
Error		146.72			
Dependent variable: Posttest-Velocity					
Model		179.64	2.90	.006	.290
SAT	1	16.52	2.40	.126	
Graph-Misc	1	18.67	2.71	.104	
Treatment	3	39.21	1.90	.139	
Sex	1	13.26	1.93	.170	
Sex x Treat.		14.24	0.69	.562	
Error		440.58			

TABLE 4-13. Sex differences on pretest and posttest scores for each treatment. Standard deviations are given in parentheses.

Treatment	n F	n M	Distance F	Distance M	Velocity F	Velocity M	Total F	Total M
				Pretest				
Standard MBL	9	9	3.00 (1.22)	3.44 (0.88)	2.56 (1.01)	3.78 (1.20)	5.56 (1.59)	7.22 (1.64)
Delayed MBL	12	7	3.17 (1.19)	3.57 (0.79)	2.58 (1.00)	3.43 (1.27)	5.75 (1.42)	7.00 (1.63)
Pencil-Paper	8	10	3.50 (0.76)	3.40 (1.07)	2.50 (1.20)	2.50 (1.35)	6.00 (1.31)	5.90 (1.91)
Test Only	7	13	3.71 (0.49)	2.92 (1.38)	2.14 (1.77)	2.85 (1.46)	5.86 (1.77)	5.77 (2.28)
				Posttest				
Standard MBL	9	9	7.44 (0.73)	7.11 (1.05)	9.33 (2.24)	12.56 (1.59)	16.78 (2.54)	19.67 (2.35)
Delayed MBL	12	7	5.42 (1.62)	6.43 (1.90)	9.25 (3.08)	10.14 (2.61)	14.67 (2.96)	16.57 (2.99)
Pencil-Paper	8	10	4.38 (2.20)	5.80 (1.99)	9.62 (2.56)	10.00 (3.23)	14.00 (3.34)	15.80 (3.97)
Test Only	7	13	5.57 (1.13)	4.62 (2.02)	8.43 (3.05)	8.92 (2.90)	14.00 (3.06)	13.54 (3.93)

TABLE 4-14. Analysis of covariance to examine treatment effects for posttest scores separated by sex of the student.

Variable	df	TypeI SS	F	p	r^2
Dependent variable: Posttest--Females					
Model	5	102.78	2.75	.037	.315
SAT	1	27.99	3.75	.062	
Pretest	1	23.21	3.11	.088	
Treatment	3	51.15	2.28	.099	
Error	30	223.97			
Dependent variable: Posttest--Males					
Model	5	391.41	11.06	<.001	.633
SAT	1	62.12	8.76	.006	
Pretest	1	50.91	7.19	.012	
Treatment	3	81.56	3.84	.019	
Error	32	226.59			
Dependent variable: Distance Posttest--Females					
Model	5	48.97	4.31	.005	.418
SAT	1	4.95	2.17	.151	
Pretest	1	0.92	0.40	.530	
Treatment	3	42.15	6.18	.002	
Error	30	68.25			
Dependent variable: Distance Posttest--Males					
Model	5	75.76	6.50	<.001	.504
SAT	1	8.47	3.63	.066	
Pretest	1	15.32	6.58	.015	
Treatment	3	15.59	2.23	.104	
Error	32	74.55			
Dependent variable: Velocity Posttest--Females					
Model	5	33.16	0.83	.540	.121
SAT	1	2.30	0.29	.596	
Graph-Misc	1	8.70	1.09	.306	
Treatment	3	11.66	0.48	.695	
Error	30	240.39			
Dependent variable: Velocity Posttest--Males					
Model	5	128.51	4.20	.005	.396
SAT	1	18.51	3.03	.092	
Graph-Misc	1	9.95	1.63	.211	
Treatment	3	32.73	1.78	.170	
Error	32	195.81			

significant (p>.10) for either subtest, but it was sufficiently consistent to indicate a significant treatment effect (p=.019) for the total kinematics posttest (combination of the two subtests). This was significant, however, only between the standard-MBL and the test-only treatments which were extreme on the posttest measure.

Considering the very short treatment period (one class period), these results are encouraging. They suggest that this type of experience may help females develop graphing skills, although females may benefit from additional practice with graphs of properties that are neither too complex nor too abstract. At the same time, males were also developing graphing skills with the more abstract concept of velocity, and therefore, in the short term, the MBL experience may have widened the sex differences in graphing. Rather than provide answers, these results define questions to be addressed by future research.

Summary

In spite of the very short treatment period, significant treatment effects were detected, using both the individual students and the treatment groups as units of analysis. After controlling for significant covariates, students in the standard-MBL treatment made fewer errors on the posttest than did students in either the delayed-MBL or the pencil-and-paper treatments. This was almost entirely due to differences on the distance subtest. This indicates that most of the effectiveness of MBL in developing cognitive links between an event and the graph of the event is associated with the immediacy of real-time

graph display. Some caution must be exercised in interpreting these results because of the statistical bias produced by the nonequivalence of treatment groups on SAT scores.

Item analysis of the responses on the posttest indicated that students made errors with (a) understanding the concepts, (b) linking the scientific description of the concepts with the real-world event, and (c) linking the concepts with the corresponding graph. The reduction in errors by students in the standard-MBL treatment appears to have been restricted to learning conventions of graphing representation and something about the concept of distance. There was no indication of their having learned about the concept of velocity.

Females consistently made more errors with items of velocity graphs than did males, even after allowing for their lower SAT and development scores. The standard-MBL experience seems to have helped the females to improve their performance, but only with distance graphs. In overall performance (both distance and velocity graphs) males may have benefited more than the females.

CHAPTER V

GRAPHING SKILLS--RESULTS AND DISCUSSION

Introduction

In the previous chapter, the results suggested that most of the improvement in posttest performance with the standard-MBL students was with (a) graphing conventions and (b) the distance concept. Throughout this study there has been consistent evidence of much greater difficulty interpreting graphs of velocity than graphs of distance. In this chapter, performance on tasks requiring graph comprehension and graph construction skills are examined to further our understanding of what kinds of difficulties students have with graphs, and what kinds of students have problems.

Performance in comprehending or interpreting graphs depends on two interwoven factors, understanding of the conceptual content of the graph and competence with graphs as a system of representing concepts. In this chapter, items in the graph test are analysed to reveal characteristics of the conceptual content of the graphs that influence comprehension of the graphs and the graphs constructed by students (from data on trials of bouncing a ball) are examined for evidence of how much students comprehend the conventions and functions of graphs as a system for representing data. This information is related to attributes of the

students and their pretreatment ratings of attitudes to graphs. Particular attention to gender differences continues.

Interpretation of Graphs

Students with more science courses had better scores on the graph test (r=.346, p=.001), but these scores were not correlated with the number of mathematics courses (r=.172, p=.101) (N=93, Table 4-2). This is interesting considering that most of the explicit instruction in graphing skills takes place within mathematics courses. As noted in Chapter III, however, students may not have been consistent in their classification of math and science courses and these covariates may not be reliable.

Correlations of graph test scores with SAT and development were both highly significant (p<.001) (Figures 5-1 and 5-2). McKenzie and Padilla (1984) also found correlations between development and graphing skills, both measured by tests different from ones used in this study. Although SAT and development scores were highly correlated with each other, they were both included by stepwise regression as covariates in the general graph test score. After partitioning variance explained by SAT, development scores still accounted for 9% of the remaining variance.

Graph Test = 1.874 + 0.0075 x SAT + 0.228 x DEV [r^2=.366, p<.001]

[SAT partial r^2=.280, p<.001; DEV partial r^2=.086, p<.001]

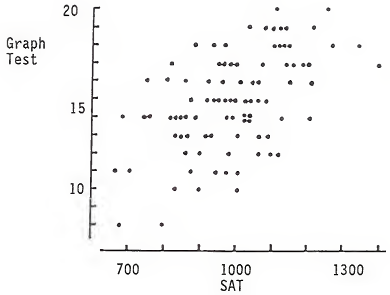


FIGURE 5-1. Correlation between students' scores on the graph test and SAT (r^2=.529, p=.0001, n=90).

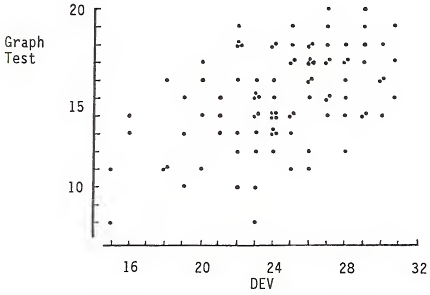


FIGURE 5-2. Correlation between students' scores on the graph test and on the test of reasoning and development (DEV) (r^2=.468, p=.0001, n=93).

Students' performance on the graph test [mean 14.96 (75%), s.d. 2.70] before treatment did not indicate at first glance that there was a serious problem with their ability to interpret graphs, although two females had scores of only 40%, representing little more than chance probability. Mokros and Tinker (in press) noted that, for middle-school students, the easiest items on a test of interpreting graphs were those for which the mental image of the event described resembled the line on the correct graph, and the most difficult items were those where the mental image was discrepant from the correct graph. Unfortunately all of their difficult items involved distance and velocity graphs.

The miscellaneous-graph subtest provides examples of items dealing with properties other than velocity. The items differ in conceptual difficulty. An item analysis of performance on this subtest (Table 5-1) showed a clear trend of graphs of simple, measurable properties (e.g., depth, height, water level) being considerably easier to interpret than graphs of more complex or abstract properties (e.g., total income, area, gas consumption). Item numbers in this table provide item identification only: They do not imply a sequenced order of presentation. The six items dealing with simple properties, such as height and depth, each had an associated mean error rate of less than 10%. In contrast, the three items with graphs of more complex properties for the dependent variable (questions 12, 9, 11) had much higher error rates--39%, 59%, and 73%.

Some errors on the most difficult item (question 11) may have resulted from some ambiguity in the item. The graph was described as "gasoline use during travel on flat land", implying a constant rate, but

TABLE 5-1. Mean error rate (MER %) of miscellaneous-graph subtest items.

Item	MER %	Description of Event	Axes
6	3	Growth of a tree over several years	height-time
7	4	Sand castle being washed away by the waves	height-time
10	5	Water level of a river through seasonal rains and dry spells	water level-time
8	7	Water draining from a bath tub	depth-time
18	7	Yo-yo moving up and down rhythmically	height-time
17	9	A seat on a ferris wheel	height-time
13	22	A burning candle	length-time
12	39	Income from a job with an hourly wage	total income-hours worked
9	59	Area of a square	area-length of side
11	73	Gasoline use during travel on flat land	gas used-distance travelled

the label on the graph was "gas used", a property which would increase with time. Similarly, the remaining item (mean error rate of 22%) involved "a burning candle." This item may also have generated some confusion because the ordinate was simply labelled "length." It is not explicitly stated that the length applies to the candle rather than the flame or period of candlelight. The alternative graph most frequently selected showed a constant length for some time, then an instantaneous drop to zero length. Thus, two of the four most difficult items may have been ambiguous.

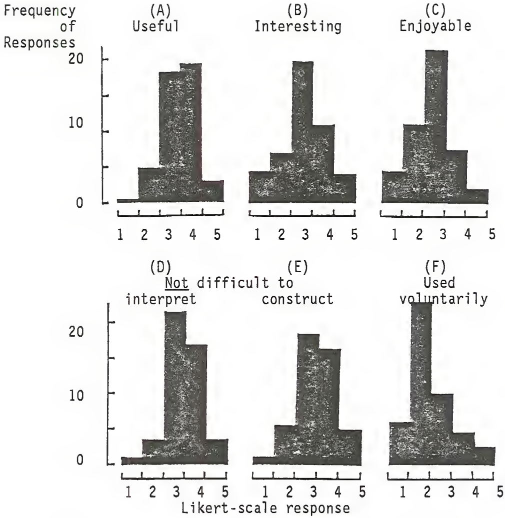
In the speed-graph subtest, the three most difficult questions involved events where the pictorial representation (spatial attributes of the mental image of the event) was strong and discrepant from the

correct graph. This is consistent with findings by Mokros and Tinker (in press). The events were "Climbing a hill and sledding down," "Bicycle up a steep hill and down the other side," and " A wagon rolls down a hill into a tree."

One pair of items on the speed-graph subtest is of particular interest. Question 1 was "A car stops at a stop sign then moves on." Question 20 was "A car slowing down, then speeding up." The sole difference between the correct graphs for these two items, was that the speed drops to zero in question 1, but it remains positive (just above the abscissa) in question 20. The general shapes of the correct graphs were the same and, by good fortune, the alternative graphs offered were the same in the two items. However, the error rate for question 1 (21%) was three time higher than the error rate for question 20 (7%). All the possible contributing factors (order of presentation, clarity of verbal description) would seem to favor question 1, not the reverse. It is possible that question 1 was perceived as the equivalent of two graphs combined, whereas question 20 was seen as a single continuous movement.

Attitudes to Graphs

The pretreatment questionnaire (Appendix F) provided information regarding students' attitudes about graphs and graphing. Responses to each item are shown in the frequency distribution in Figure 5-3. All Likert-scale items were scored with 5 being the most positive rating. As discussed in the methods chapter, a probability of .10 was used in analyzing the statistical significance of these attitude data.



Graph	Abbreviated Item Description	Mean	s.d.
A	Graphs are useful	3.43	0.80
B	Graphs are interesting	2.91	1.04
C	Enjoy working with graphs	2.84	0.95
D	Not difficult to understand	3.38	0.81
E	Not difficult to construct	3.43	0.88
F	Graphs used voluntarily	2.45	1.01

FIGURE 5-3. Frequency distribution of student's pretreatment responses to a questionnaire assessing their attitudes to graphs. Responses are on a Likert scale, with 5 being the most positive response.

Although students generally considered graphs useful (mean 3.43, s.d. 0.80), their ratings of interest (mean 2.91, s.d. 1.04) and enjoyment (mean 2.84, s.d. 0.95) were significantly lower ($p<.05$) than their ratings of utility. Students reported that they do not often use graphs voluntarily (mean 2.45, s.d. 1.01). Only 14 out of the total of 93 students reported that they use them "often" or "nearly always."

Females rated graphs as more difficult ($p<.05$) and less enjoyable ($p<.10$) than did males (Table 5-2). Results presented in Table 4-11 indicated that, although females had more difficulty, most of this difficulty was accounted for by differences in ability and development.

Based on Pearson correlation coefficients between responses on each item, items were grouped into two categories, interest (items 11, 12, and 15) and difficulty (items 13 and 14). Correlation coefficients between these two categories were not significant ($p>.20$), but correlation coefficients among items within each category were highly

TABLE 5-2. Sex differences in attitudes to graphs in science courses, scored by responses on Likert-scale items, with 5 being the most positive.

Abbreviated description	Females (n=45) Mean	s.d.	Males (n=47) Mean	s.d.	sig.
Graphs are useful	3.33	(0.80)	3.53	(0.80)	
Graphs are interesting	2.87	(0.92)	2.96	(1.16)	
Graphs are enjoyable	2.64	(0.83)	3.02	(1.03)	*
Not difficult to understand	3.20	(0.89)	3.55	(0.69)	**
Not difficult to construct	3.22	(0.90)	3.64	(0.82)	**
Used voluntarily	2.45	(1.04)	2.45	(1.00)	

correlated ($p<.001$). This indicates that each category represents a separate measure of affect. Although ratings of usefulness were significantly correlated with both interest ($r=.465$, $p<.001$) and enjoyment ($r=.353$, $p<.001$), Rowe and Hurd (1966) found, in a study of the Biological Science Curriculum Study, that usefulness of an activity was a better predictor of application to another situation than was interest.

Figure 5-4 shows why the correlation between these two categories of items (interest and difficulty) was not significant. Many students believed that they had few difficulties with graphs and yet had little interest in them. This is consistent with the literature on motivation via curiosity, which suggests that people have little interest in attending to things that they believe they already know (e.g., Berlyne, 1957, 1978).

Students' performance in graphing skills may reflect their attitudes, whether they perceive graphs as interesting or difficult. Because these two categories of attitudes were not correlated, it is interesting to examine whether either of them was statistically correlated with graphing performance.

Correlation coefficients among these categories of attitudes to graphs (interest and difficulty) and other covariates and performance measures are provided in Table 5-3. Perceptions of interest and difficulty with graphs were not significantly correlated with either measure of aptitude (SAT or development), or with scores on the kinematics pretest and posttest. Also, performance on the graph test was not significantly correlated with ratings of interest in graphs

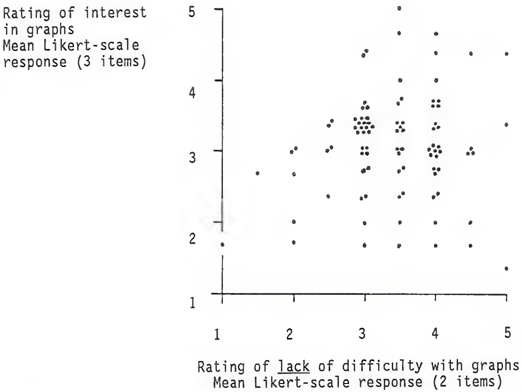


FIGURE 5-4. Correlation between student's pretreatment rating of their interest in graphs (3 items) with their perceived lack of difficulty (2 items) in using graphs. Responses are on a Likert scale, with 5 being the most positive response.

TABLE 5-3. Pearson correlation coefficients between attitudes to graphs and covariate and performance measures. Probabilities of _r_>0 are given in parentheses.

Category of attitudes	SAT	DEV	Graph	Pretest	Posttest
Graphs are interesting	.076 (.477)	.013 (.904)	.094 (.371)	.071 (.503)	-.013 (.909)
Graphs are not difficult	.131 (.218)	.121 (.252)	.254 (.014)	.081 (.443)	.177 (.105)

(r=.094, p=.37). Students with higher ratings of difficulty with graphs had lower scores on the graph test (r=.254, p=.014). Although this seems to indicate that their perceptions were reasonably accurate (i.e., appropriate metacognition of their graphing comprehension skills), it is by no means clear whether the perception of difficulty was the cause or effect of poor performance.

Construction of Graphs

Tests used to assess comprehension of graphs have generally asked students to select the appropriate graph, or, given a graph or series of graphs, to interpret the relevant attributes or read specific information from them. Like the pretest, posttest, and graph test used in this research, they are generally multiple-choice. Although this can provide information about the ability to read from and interpret graphs, it does not provide insight into the ways that students think about and use graphs in a science context. In other words, it demonstrates whether students know the appropriate algorithms to use, rather than whether they have an appropriate graph schema, or whether they understand the meaning and utility of graphs in a given context.

In tasks requiring graph comprehension, the graph schema specifies how to (a) translate information from a visual image to a conceptual message, (b) translate a conceptual question into a search strategy, and (c) recognize which type of graph, complete with pertinent grammar, is appropriate for a given context. From the results reported so far (e.g., performance on the graph test, Table 4-9), it appears that

students in this study had the requisite algorithms for the first of these (a), and MBL experiences seemed to have provided experience and practice with the second (b) (see Chapter VI, behavioral observations). The ability of students to invoke and implement an appropriate graph should reveal some information about the third of these processes.

In the posttreatment testing session, in an attempt to expose the understanding of the conventions and syntax of Cartesian graphs, students were asked to construct a graph using data provided from five trials of bouncing a ball. Data included the height from which the ball was dropped, and height to which the ball bounced. Students were asked to pay attention to specific features adapted from the Test of Graphing Skills developed by McKenzie and Padilla (1986). These features (detailed in Padilla et al., 1985) were scaling axes, assigning variables to axes, plotting points, and a best-fit line (described in the test questions as "showing the trend in the set of data points"), and labelling axes (not included in Padilla et al.'s characterization).

Students had a great deal of uncertainty and difficulty in constructing this graph. Difficulties ranged from minor problems through clear omissions and errors to primitive or nonexistent understanding of representation of data on a Cartesian graph. Some problems are minor, but others are major. The frequencies of occurrence of the different kinds of errors or problems are shown in Table 5-4.

The minor problems described here included incomplete labels, transposing the dependent and independent variables (reversing axes), unbalanced scales, and connecting the points rather than drawing a straight best-fit line. Not everyone would agree that these are indeed

TABLE 5-4. Frequency of various deficiencies and errors in constructing a graph (84 students).

Graph feature	Type of deficiency	Occurrence Rate (%)
Axes	Correct:	35
	Error: Axes reversed	48
	Graph as picture	13
	Wrong variable	11
Labels	Correct:	35
	Error: Absent	10
	Incomplete	45
	No units	25
Scale	Correct:	51
	Error: Irregular	13
	Unbalanced	30
Points	Correct:	75
	Error: Wrong	13
	Missing	5
Best-fit-line	Correct:	6
	Error: Point-to-point	63
	Histogram	6
	None	6

errors at all. At a recent conference, several researchers and teachers discussed some of these issues. Priscilla Laws (personal communication, May, 1986) correctly pointed out that the appropriateness of drawing a straight best-fit line depends on the precision of estimating each point. Some thought that it was irrelevant which variables were assigned to which axes. Others suggested that, because students are often required to connect data from a time series from one point to the

next, it is unreasonable to expect them to draw a straight best-fit line. It is also acknowledged that, technically, unbalanced scales are not incorrect. Simply, they fail to optimize the spatial attributes of the graph for revealing the pertinent relationship. It is not only essential to use attributes of graphs correctly, it is also important to avoid misusing them (Jones, 1985).

These kinds of debatable interpretations and criteria are perhaps a considerable part of the reason why such tests are seldom used in research. With many of the "problems" in constructed graphs (Table 5-4), we do not know either what the errors signify, or even whether they should rightfully be regarded as errors. Nevertheless, used cautiously, they can provide some valuable information, and for this purpose, it is more useful to be overrigorous than overlenient in the characterization of errors. Whether these objections and/or qualifications to the listing of problems in Table 5-4 are reasonable, they are largely irrelevant to the purpose of making the categorization, which was to reveal the quality of students' graph schemas.

The adequacy of graphs constructed is presented in Table 5-5. Although no student produced a graph that I considered to be entirely satisfactory, seven students (out of the 84 who completed this question) were only slightly short of this, with a single minor inadequacy. They appeared to have a thorough understanding of the grammar of graphs. Most constructed graphs with several relatively minor errors. Whether the deficiencies were the result of carelessness or lack of understanding, the sheer number of them points to a substantive lack of confidence and/or competence with graphing techniques and conventions.

TABLE 5-5. Levels of adequacy in graphs constructed.

Level of adequacy	Number of students
Completely satisfactory	0
Single minor inadequacy	7
Several minor deficiencies	31
Major errors or omissions	28
Primitive or preCartesian	18

A fifth of the students produced some kind of "primitive", preCartesian graph, in the sense that they utilized only one axis to represent both continuous variables (height of drop, height of bounce). Representative graphs from most categories of preCartesian graphs are shown in Figure 5-5. The frequencies of graphs in each category can be seen in Table 5-6. Of the graphs in Figure 5-5, the bar chart ordered by height (graph F) is the only one which utilizes (albeit inadequately) continuous variables on both axes, and these graphs were not included among the primitive graphs for further examination. Graph B (Figure 5-5) uses a spatial organization of data, but they are only scaled in one dimension (i.e., the data points do not have Cartesian coordinates). The categories in Table 5-5 represent a continuum from graph-as-picture concept to a bar chart, or a progression in the degree of complexity, abstraction, and conceptual sophistication that parallels the order in which graphic representation is introduced through the school curriculum.

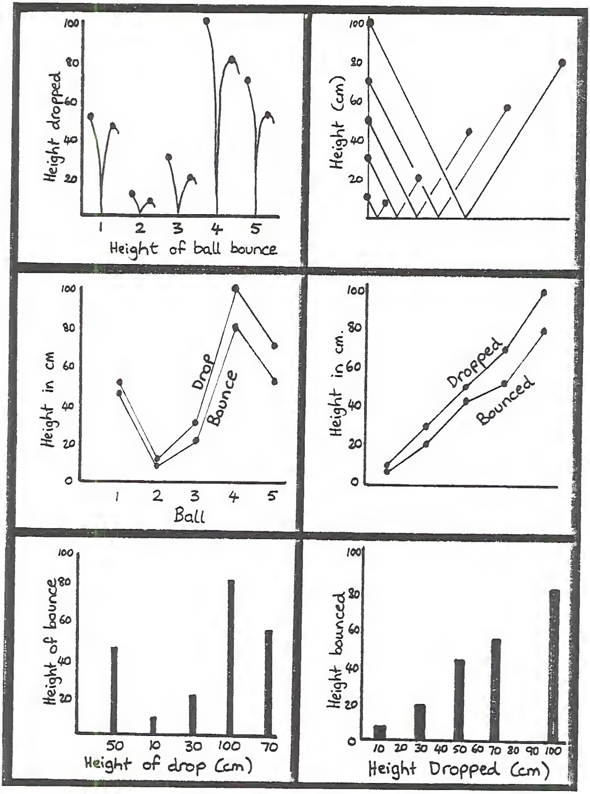


FIGURE 5-5. Examples of preCartesian graphs constructed by students to represent data from five trials of bouncing a ball.

TABLE 5-6. Types of preCartesian graphs constructed. The first five of these categories involved the use of a single continuous-scale axis.

Categories of Graphs		Fig. 5-5	Frequency
Graph-as-picture	ordered by trial	(a)	9
Graph-as-picture	ordered by height	(b)	2
Two lines/single axis	ordered by trial	(c)	2
Two lines/single axis	ordered by height	(d)	2*
Bar chart	ordered by trial	(e)	2
Bar chart	ordered by height	(f)	4
Two axes but no grid			1

* It is likely that one of these students copied from the other.

Half of the primitive graphs illustrate the graph-as-picture problem that has been described by other researchers (Clement, 1985; Clement et al., 1986; Mokros & Tinker, in press). For the other graphs, the conceptual understanding of graphing skills is ambiguous. For instance, in graph D (Figure 5-5), the students distinguished between the two variables, but did not follow the conventions for showing the relationship between them. Their intention is not clear. Similar interpretive problems are presented by the graphs where the data are ordered by trial sequence. Most of the graphs with which students have experience in science courses have time as the independent variable, so time becomes a dominant clue in constructing graphs. It is possible that students who constructed graphs where the data were ordered by trial number, perceived the implicit trial number as a pseudo-time measure. They failed to appreciate that graphing data is not a simple transcription process, but it also involves organization of data to reveal pertinent relationships.

Students who constructed primitive graphs had a general graph schema that was either inadequate or nonfunctional. In six years of learning and using graphs, most of these students had learned algorithms for coping with graphs, but they did not have a clear understanding of what a graph is or what its function is. Most of these students were about to graduate from high school and go on to college, where lecturers would assume that, among other entry skills, students would have adequate graphing skills (Arons, 1979). And these students would not.

Because females reported that they had more difficulty with both graphs and the MBL activity than did males, information from the 18 students described above (who used preCartesian graphs) was examined to see what personal attributes (including sex) may have been related to their problems with a graph schema. The data, separated by sex, are compared with data for the total population (Table 5-7). They were not separated by treatment because the numbers were so few.

In general, there are two surprises: (a) that two thirds of the students with graph schema problems were males, and (b) that the males were similar to the average of all students on measures of SAT, development, and graph test. Indeed, one of the males had a SAT score of 1210. The poor graph schemas of the male students do not appear to have been the result of low aptitude. In contrast, all of the females were in the lowest quarter for SAT scores, five of the six were in the lowest quarter for scores on the graph test, and their mean development score was also low. For these females, the inadequate graph schema may reflect comparatively low general ability.

TABLE 5-7. Descriptive information about students who constructed preCartesian graphs.

	All Students	Students with Graph Schema Problems	
		Male	Female
	(*N*=93)	(*n*=12)	(*n*=6)
Covariates			
SAT	990	928	762
DEV	24.4	24.5	20.8
Graph test	15.0	14.2	11.0
Rated Attitudes to Graphs			
Useful	3.4	3.1	3.0
Interesting	2.9	2.67	2.75
Not difficult	3.4	3.42	3.08
Kinematics Performance			
Pretest	6.11	5.67	5.50
Distance Posttest	5.77	5.36	5.67
Velocity Posttest	9.71	9.36	8.50

Sex differences were apparent when students with major inadequacies in their graph schemas were compared with the total population for (a) their rated attitudes to graphs, and (b) their performance on pretest and posttest. Both males and females in this group tended to rate graphs as less useful than did the rest of the students. Females indicated more difficulty with graphs, and males indicated slightly lower interest in graphs. On performance measures, both males and females had lower scores for the pretest and velocity section of the

posttest, but this was not striking. The general pattern that emerges from these data is that, although the females may have had poor graph schemas because of low ability, the performance of the whole group was sufficiently poor as to indicate that previous instruction may have been inadequate for students to develop a functional graph schema. Because of the very low sample numbers of students being compared, these data can be examined only for indicators, not for statistical differences.

MBL and Graphing Skills

To see whether there was any transfer of graphing skills to other types of graphs, an analysis of covariance was performed on the graph test, the miscellaneous-graph subtest, and the speed-graph subtest, which were administered during the posttreatment testing session. After controlling for significant covariates determined by stepwise regression (mainly the same test used as a pretest), the analyses failed to indicate a significant ($p<.05$) treatment effect. This is not surprising because the mean change in score from graph pretest to graph posttest was an increase of only 0.35 items (i.e., <2% improvement).

These results do not provide evidence that any improvement in graphing skills from the MBL experience transferred to more general tasks of interpreting graphs. However, the null result should not be overinterpreted because of the extremely short treatment period used in this research. Linn et al. (in press) have provided some evidence for transfer of graphing skills from an extended MBL experience with a

temperature unit to performance in interpreting graphs of motion within a similar testing situation.

Recently, two separate studies of extended experience with MBL (Linn et al., in press; Mokros & Tinker, in press) have shown substantial gains in graphing skills by middle-school students. Nevertheless, in both studies, a core of about one third of the students retained fundamental confusions or misconceptions. That certainly indicates some kind of resistance. It is not clear whether the remaining problems derive from an inadequate graph schema, an inadequate conceptual understanding of the concept being graphed, or an inadequate command of cognitive skills for determining what variables are relevant to the task. This issue warrants further investigation.

An anecdote will illustrate the complexity of the problem. During pilot studies, I spent an extended period observing a tenth grade physics class as they worked through the kinematics section of their course. These students completed the pretest used in this dissertation research, and many of them exhibited slope/height errors in their selection of answers. Before reviewing the correct answers with them, I had students come to the front of the class and act out being either a distance detector or a velocity detector. As I moved towards and away from them, I instructed them "to show with their hands" what happened to the property they were detecting. These students had no hesitation or difficulty in gesticulating the trend of the graph. They did not seem to have difficulty with the representation of the event, but they needed some instructional guidance, which was restricted to questions such as "Is the distance (or speed) increasing, decreasing, or staying the

same?" to help them to focus on the relevant concept and discrimate from the irrelevant ones. They had difficulty focusing on a single relevant concept, but once they focused on this concept, they had no difficulty in demonstrating the appropriate trend. This incident, as well as others observed during the study, suggested that perhaps the difficulty is less with either the concepts or the representation, but instead lies with students' command of mathemagenic skills--the ability to identify relevant concepts and isolate them from the mire of other concepts and cues in the contextual environment.

Summary

Although students in this study had few errors interpreting graphs of simple physical properties, they had much greater difficulty with graphs of more abstract or complex entities or where the mental image of the event was discrepant with the graphic representation of it. The error rate for such items exceeded one third.

On the pretreatment questionnaire, students reported that they generally thought graphs were useful, and not difficult, but not very interesting either. Females rated graphs as more difficult than did males, and they also enjoyed using graphs less than did males. Performance on the graph test was correlated with perceived lack of difficulty with graphs, but not with interest.

One out of five students was unable to produce a functionally adequate graph to represent given raw data. This is seen to be a serious deficiency in basic skills, reflecting an inadequate graph

schema. Inadequate or undeveloped graph schemas may underlie problems with graphs that persist after extended experience with MBL activities.

CHAPTER VI

BEHAVIOR AND ATTITUDES

Introduction

The analysis of the quantitative posttest results indicated that there were improvements in performance of students from standard-MBL groups compared with both delayed-MBL and pencil-and-paper groups, even though the intervention period was very short. Most of the difference appears to have been related to the real-time graphing feature of the standard-MBL treatment. Although this difference in performance was significant only for the distance subtest, it is possible that, if there were a larger number of students in each treatment and/or a longer intervention period, similar significant differences would be apparent for the velocity-time subtest.

One of the questions that arises from these results is whether this improvement in performance resulted from differences in cognitive factors, in motivational factors, or in both. It is important to examine student behavior and attitudes during the treatment, and to pay particular attention to whether any behavioral differences can be related or attributed to the real-time versus delayed-time graphing. To guide this investigation, there were a number of indicators that could be used--entries on student worksheets, questionnaires of posttreatment

attitudes to the activities, posttreatment comments volunteered by the students, and observations of students' behavior made both during the activities and later from transcripts and videotapes. Of these sources of information, the only quantitative data came from the posttreatment questionnaire.

Questionnaire--Attitudes to the Experimental Activities

Students' responses to Likert-scale items (Appendix F) about their attitudes to the activity are shown in Table 6-1 and Figure 6-1. Responses were rated with the most positive response being scored as 5, and the least positive response being scored as 1. The question asking whether they were interested in further activities was not clearly worded, so it has been excluded from consideration. Of the remaining eight items, all but the last directly address perceptions and attitudes to the activity itself. The last item, "How helpful were the other members in your group?," addresses perceptions of the social environment of the activity.

Very few students reported extreme ratings of the difficulty level of the activities. Only two students reported that the activities were "not at all" difficult, and only two students reported the activities were "very much" confusing. Students in all three treatments found the activities interesting and useful. They enjoyed the activities, and thought they helped them to understand both motion and graphing.

Based on Pearson correlation coefficients between responses on each item and on construct similarity, items were grouped into

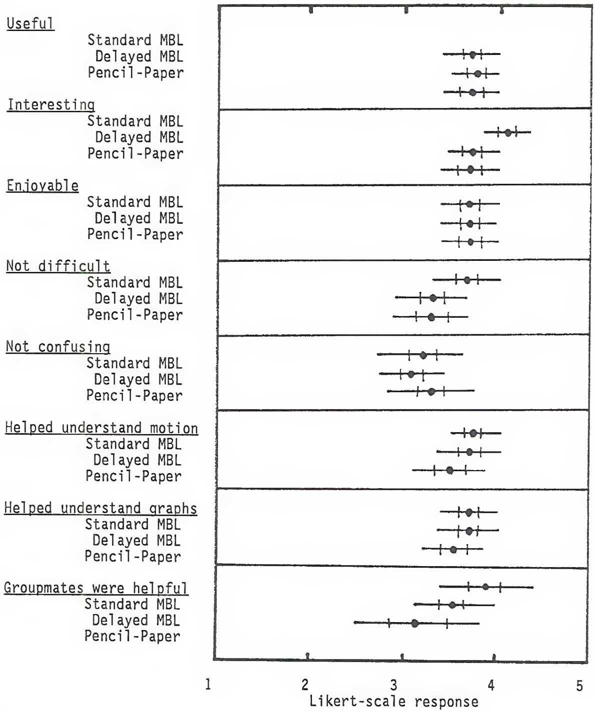


FIGURE 6-1. Posttreatment ratings of attitudes to the experimental activity. Responses are on a Likert scale, with 5 being the most positive response. Mean, 50%, and 95% confidence limits are shown.

TABLE 6-1. Students' responses to Likert-scale items in the posttreatment questionnaire to assess attitudes to treatment activity (mean and standard deviation). The full text of the items is provided in Appendix F.

Item description	Standard MBL		Delayed MBL		Pencil-Paper	
	Mean	s.d.	Mean	s.d.	Mean	s.d.
Useful?	3.67	0.658	3.74	0.619	3.67	0.594
Interesting?	4.14	0.578	3.74	0.689	3.72	0.575
Enjoyed?	3.90	0.768	3.70	0.765	3.72	0.575
Not difficult?	3.71	0.845	3.30	0.926	3.33	0.767
Not confusing?	3.24	1.044	3.09	0.848	3.28	0.958
Help understand motion?	3.76	0.625	3.70	0.822	3.50	0.786
Help understand graphs?	3.71	0.717	3.61	0.839	3.56	0.705
Groupmates helpful?	3.90	1.136	3.57	0.992	3.17	1.383

categories that paralleled categories for the pretreatment questionnaire for attitudes to graphs. The categories were interest (items 21, 23, 25), difficulty (items 22 and 24), and helpful to understanding (items 26 and 27). Correlations between items within each category were highly correlated with each other ($p<.01$) with the exception that ratings of the usefulness of the activity were not significantly correlated with rating of how interesting it was ($r=.189$, $p=.141$). Correlation coefficients between these categories indicated that the ratings of interest were not significantly related to the ratings of difficulty ($r=.153$, $p=.24$), but they were significantly related to ratings of how much they helped understanding ($r=.482$; $p=.0001$).

Pairwise comparisons of responses to individual items and categories indicated that the students in the standard-MBL rated their overall activity (all items combined), and particularly how interesting was the activity as significantly (p<.10) more positive than did students in both delayed-MBL and pencil-and-paper groups. The same pattern was apparent where students rated how helpful were their groupmates, although in this case, because of the increased variability among the responses, the difference was significant only between the standard-MBL and the pencil-and-paper treatments.

This item, how helpful were the groupmates, is particularly interesting because it measures social environment rather than the activity itself. Moderating factors such as this provide the rationale for statistical analysis using the group (or other higher organizational levels) as the unit of analysis rather than the individual. Although students in the pencil-and-paper group were instructed to work as a group, and they generally did so, their activities did not necessitate that they work together. It is not surprising then that these pencil-and-paper students were neutral in their rating of the helpfulness of their group members.

Students in the two MBL treatments generally thought that their group members were indeed helpful, with the students in the standard-MBL treatment tending to respond more positively than students in the delayed-MBL treatment. Because the treatment with the highest posttest performance also gave the highest ratings to groupmates for helpfulness, it seems likely that the higher ratings were the result of either improved learning itself, or, even more likely, the result of some other

variable, such as improved sense of competence, efficacy, or motivation, which was itself correlated with the higher performance.

This line of investigation can be extended to individual students (rather than students aggregated by treatment) by examining the correlation between individual scores on posttest and their ratings of how helpful were their groupmembers within each treatment (Fig. 6-2). With students in the pencil-and-paper treatment, there was a significant negative linear regression between these measures (r^2=.292, p=.021, n=18). This negative correlation is not apparent with students in either of the two MBL treatments. The difference seems to be due to higher ratings of peer helpfulness by the higher ability students in MBL groups (Figure 6-2). The MBL experience may prove effective in promoting positive social interaction.

The activities were based on graphing as the system for representation of the physical properties--distance, velocity, and time. Because graphing is so prominent in the activities, it is interesting to note that the correlations between students' pretreatment ratings of interest in graphs and their posttreatment pretreatment ratings of interest in the activity were not significant for any of the treatments.

Focusing on students who indicated a strong dislike of graphs (average rating of no more than 2 for interest and enjoyment with graphs) (Fig. 6-3), all but one of those students who were in the two MBL treatments (i.e., 10 out of 11 such students) responded to the activities with strong positive attitudes (average rating of at least 4 on items about interest and enjoyment with the activity). However of the four such students in the pencil-and-paper treatment, only two

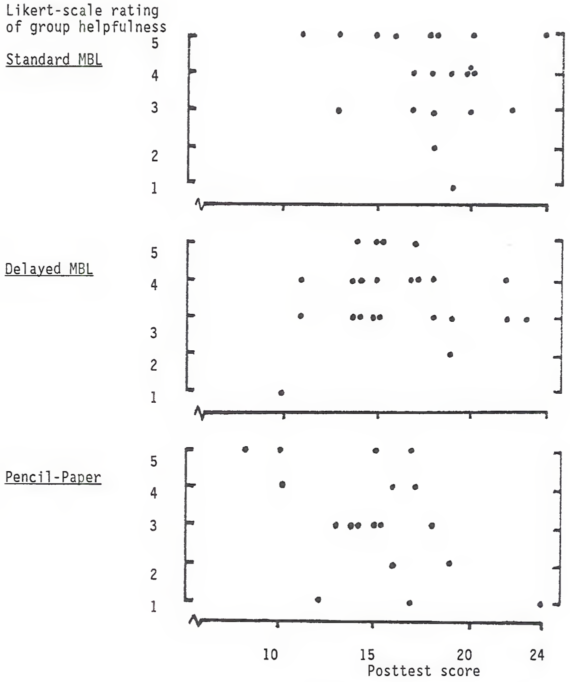


FIGURE 6-2. Correlation between posttreatment rating of helpfulness of group members (Likert-scale items with 5 being the most positive) and posttest scores.

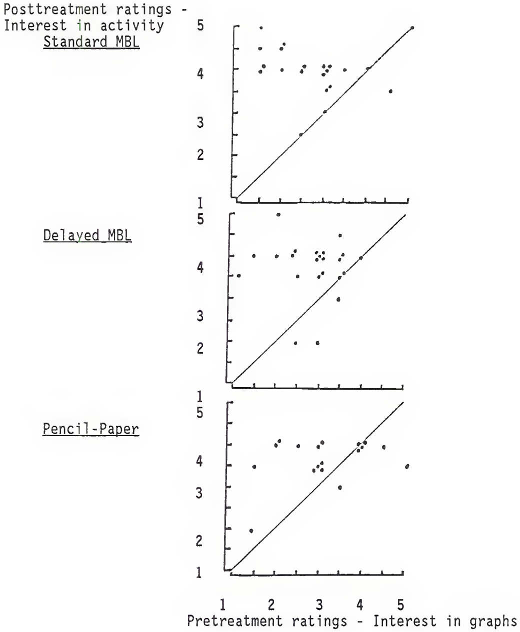


FIGURE 6-3. Correlation between students' mean pretreatment ratings of interest in graphs and posttreatment ratings of interest in the experimental activity. Ratings are on a Likert scale, with 5 being the most positive.

indicated strong positive attitudes, while one indicated a strong negative attitude to the activities. It appears that MBL experience provided an intrinsically interesting and motivating environment in which to work with and learn about graphs.

In Table 6-2, ratings of attitudes within each category are correlated with covariates and performance measures for each treatment. For all treatments, performance on the graph test was significantly correlated with the perceived lack of difficulty of the activity (p<.10), and, for the pencil-and-paper treatment, with ratings in all categories of attitudes to the activity (p<.10). These data reflect the strong emphasis on graphing skills in all activities. It is interesting to note that in no case was interest in graphs significantly related to interest in the activity (p>.20), but that, for both MBL treatments, perceived lack of difficulty with the activity was significantly correlated with lack of difficulty with graphs (p<.05).

Performance on the posttest was not correlated with interest in the activities (p>.60), but, for the standard-MBL treatment only, it was correlated with perceived lack of difficulty with the activity (r=.430, p=.052). In general positive attitudes to the activity seemed to be related to development for the standard-MBL students, whereas they were more related to ability (SAT) for the delayed-MBL students. It is interesting to observe that the correlation between SAT and interest in the delayed-MBL activity was negative (r=-.438, p=>038). Examination of the data indicated that this result was largely produced by two students who were extreme on SAT scores.

TABLE 6-2. Pearson correlation coefficients for categories of ratings of attitudes to the treatment activities. Probability of r>0 are given in parentheses.

Category of Attitude items	SAT	DEV	Graph Test	Pretest	Posttest	Graph Int.	Graph Diff.
			Standard-MBL Treatment				
Interesting (3 items)	-.188 (.416)	.243 (.289)	.053 (.821)	.054 (.815)	.005 (.984)	-.111 (.631)	.341 (.131)
Not Difficult (2 items)	.290 (.203)	.575 (.006)	.378 (.091)	.317 (.162)	.430 (.052)	-.027 (.904)	.469 (.032)
Total (7 items)	.039 (.867)	.513 (.017)	.125 (.591)	.079 (.732)	.140 (.544)	.018 (.939)	.358 (.111)
			Delayed-MBL Treatment				
Interesting (3 items)	-.438 (.036)	-.177 (.419)	-.022 (.920)	-.321 (.136)	-.101 (.647)	-.115 (.600)	-.045 (.839)
Not Difficult (2 items)	.434 (.038)	.175 (.424)	.465 (.025)	.073 (.740)	.347 (.105)	.265 (.222)	.677 (.001)
Total (7 items)	-.203 (.354)	-.087 (.695)	.193 (.377)	-.241 (.268)	.006 (.978)	.143 (.516)	.244 (.262)
			Pencil-Paper Treatment				
Interesting (3 items)	.193 (.443)	.054 (.830)	.683 (.002)	.112 (.657)	-.082 (.746)	.308 (.213)	.535 (.022)
Not Difficult (2 items)	.527 (.025)	.082 (.746)	.430 (.075)	.015 (.954)	.098 (.698)	.082 (.746)	.056 (.824)
Total (7 items)	.296 (.233)	.013 (.960)	.701 (.001)	.167 (.508)	-.025 (.920)	.268 (.282)	.466 (.051)

There were significant differences between the sexes in their attitudes to the standard-MBL activity (Table 6-3). Males rated the activity as more interesting, less difficult, and less confusing than did females in the same treatment. Indeed, their ratings in both categories of interest and lack of difficulty were significantly higher ($p<.10$) than both sexes in other treatments.

TABLE 6-3. Sex differences in attitudes to the treatment activity, scored by responses to Likert-scale items (5 being the most positive). Standard deviations are given in parentheses.

Abbreviated Item Description	Standard MBL F n=12	Standard MBL M n=9	Delayed MBL F n=13	Delayed MBL M n=10	Pencil-Paper F n=8	Pencil-Paper M n=10
Useful	3.75 (0.45)	3.56 (0.88)	3.53 (0.66)	4.00 (0.47)	3.50 (0.76)	3.80 (0.42)
Interesting	4.00 (0.60)	4.33 (0.50)	3.85 (0.69)	3.60 (0.70)	3.75 (0.71)	3.70 (0.48)
Enjoyable	3.58 (0.79)	4.33 (0.50)	3.69 (0.75)	3.70 (0.82)	3.62 (0.74)	3.80 (0.42)
Not Difficult	3.42 (0.90)	4.11 (0.60)	3.08 (0.95)	3.60 (0.84)	3.38 (0.74)	3.30 (0.82)
Not Confusing	2.83 (1.03)	3.78 (0.83)	2.92 (0.86)	3.30 (0.82)	3.12 (1.13)	3.40 (0.84)
Helpful-motion	3.83 (0.39)	3.67 (0.87)	3.69 (0.63)	3.70 (1.06)	3.50 (0.76)	3.50 (0.85)
Helpful-graphs	3.75 (0.62)	3.67 (0.87)	3.54 (0.78)	3.70 (0.95)	3.50 (0.76)	3.60 (0.70)
Group members helpful	3.92 (1.24)	3.89 (1.05)	3.69 (1.18)	3.40 (0.70)	3.62 (1.06)	2.80 (1.55)

Females volunteered as many comments on the bottom of their posttreatment questionnaire as did males (Appendix G). However, their comments were mainly that the activity was "fun" or else they were only peripherally related to the educational content of the treatment. In contrast, comments from males were related to learning, understanding, and being challenged.

Behavioral Observations

Observations presented in this section have been amalgamated from a variety of sources, each of which provides a somewhat different perspective (triangulation of evidence). In synthesizing this information, the most striking feature was not how great were the differences among groups, and particularly between the two MBL groups, but rather how consistently the differences were revealed.

Audiotapes and videotapes of students participating in the activities (from both the experiment and the pilot study), being permanent records that could be reviewed at leisure, were rated using the observation instrument described in the methods chapter. They provided a means of assessing as objectively as possible the frequency of conversation at each level of increasing cognitive engagement. The tapes were not very effective in providing indications of positive and negative affect because the verbal cues were often ambiguous. It was difficult, if not impossible, to tell from audiotapes whether a simple comment like "You have a turn" reflects positive affect (an invitation to join the fun), negative affect (the student wants to opt out), or

simply being fair about everyone having a turn. Unfortunately, the quality of many of the tapes was too poor for transcription and instrument rating because of background noise level and acoustic properties of the classrooms.

Extensive journal notes were taken during the research and also during observation of several traditional kinematics laboratory activities. These notes were primarily concerned with subjective assessment of attitudinal and motivational indicators, particularly non-verbal cues. They focused on differences between groups in each class (i.e., treatments) in behaviors such as persistence, intensity of engagement with the activity, and the time and manner of disengagement.

Students' entries on worksheets during the treatment activities, and the comments they volunteered either on their posttreatment questionnaires or in person gave some insight into their personal perspective--their attitudes, interests, and priorities. Discussions with the teachers before, during, and after the experiments in their classroom highlighted some interesting pragmatic considerations and implementation issues. The regular teachers were not required to participate in the activity in any capacity. Some teachers used the opportunity to catch up on other work, whereas others followed the activities with considerable interest.

Behavioral observations are not hard data, in the sense that they are neither objective nor quantifiable. In the following presentation, behavioral observations are presented first for students in the pencil-and-paper groups. The behavior of standard-MBL students will then be described, and contrasted first with behavior of students in traditional

laboratory activities to highlight the specific contribution of this MBL, and then with the behavior of delayed-MBL students. In this discussion behavioral observations are separated into indicators of predominantly cognitive factors and indicators of predominantly motivational factors.

Students in Pencil-and-Paper Groups

Students in the pencil-and-paper groups were not audiotaped. Observations of them came predominantly from journal notes. Students appeared to enjoy the activity, and this is corroborated by their responses on the questionnaire. They seemed to particularly enjoy the opportunity to be a little creative in the final activity (make up a story to explain a graph). They were actively engaged with the task, paying attention appropriately to the key features of the distance and velocity graphs. Sudden realization experiences ("Aha!" experiences) were impressively frequent as they worked out such things as the velocity graph being flat for a constant-velocity event. The general impression was that these students were focusing so clearly and directly on the key points (described in the methods chapter), and there were so few opportunities for confusion, distraction, or complication that it would not have been surprising for these students to outperform the MBL students on the posttest.

Cognitive Indicators--Standard MBL

The MBL activity started with some orientation exercises that fulfilled several functions. They helped the students to get accustomed to the equipment and the graph format, and how to use them. They pointed the students to the key concepts of the lab--the effect of change in speed and direction of movement on the distance-time and velocity-time graphs. They also set the tone of the entire laboratory session in requiring active mental engagement, not just physical engagement, within an inductive instructional format.

Students seemed to approach the laboratory with a very habitual passivity. In the pilot study, when seats were not removed, some students would sit down for the entire class period. The original worksheets (obtained from TERC) asked students to sketch the graphs for each trial event. Even when I adapted this, so that students were asked to sketch graphs for two events on the same set of axes and then to compare the two lines, students would painstakingly copy every little bump and wiggle in the lines without really looking at what it meant.

To break this passivity, and to get students thinking as early in the period as possible, the worksheet used for the main experiment asked students to "describe (in words or sketch) what happens when you change direction" (or speed, etc.). This seemed to induce students to pay attention immediately to what the graph meant rather than what it looked like. The tradeoff was that it introduced some confusion and uncertainty about what exactly they were supposed to do. This did not appear to be severe or dysfunctional, and very few of the students rated

the activity as either difficult or confusing on the posttreatment questionnaire.

The computer performed the procedural functions of collecting the experimental data for each trial run, making all the necessary calculations, and displaying the data on the screen. During each trial run, one student moved towards and away from the sensor at different speeds in order to generate the experimental data, while another operated the computer. While the data were being collected and displayed, students could watch the "mover" or the screen display. During this orientation phase, the movers gradually learned to use the information displayed on the screen to guide or regulate their movements. There were often instructions from the acting "choreographer" to the effect that the mover was going "too fast" or "too slow," or such as "wait until <u>it</u> gets to x seconds." Most students learned to examine the data for pattern recognition, for identification of anomalies, and for explanations as the data were being displayed.

<u>More time for interpretation</u>

Initially most of the discussion was centered around procedures, observations, and comparisons of data. As the laboratory activity progressed, students spent less time with procedures (organizing how to collect data and doing so) and observations, and more time interpreting their results--explaining them, evaluating their data in the light of their expectations, in planning the next event to be monitored (generating and testing hypotheses), and refining concepts (Figure 6-4).

In effect, the MBL activities started close to where many conventional laboratories finish, with the students' interpretations of data from initial experiments providing the starting point for subsequent experiments and guiding the predictions and hypotheses being tested. The MBL activities appeared to ferment activity and involvement in a wide spectrum of science-process skills. Within a single class period there were multiple cycles of an experimental question leading to an experiment or trial event, which then produced more results to be explained, evaluated, and interpreted. In turn, this led to generation of a new set of expectations, hypotheses and questions. Each of these cycles reinforced students' understanding of the concept and helped to build their confidence in their understanding. Such multiple cycles of experimentation, such as alternating learning and challenge cycles (Hegarty, 1982; Venkatachalam & Rudolph, 1974), have long been advocated as effective for teaching science concepts and processes.

This spectrum of involvement with science processes is what science labs are intended to accomplish but seldom do. In many conventional laboratory activities, so much time is necessarily devoted to collecting data and performing calculations that there is seldom sufficient time and opportunity for interpreting the results of an experiment. Even where time is available, the process of interpretation is usually separated from the actual experiment by a considerable period, often overnight and sometimes longer. As is shown in Figure 6-5 (the lab report from a bright, tenth-grade, physics student in a conventional inclined-plane experiment), the interpretation may raise questions, but seldom are these questions pursued.

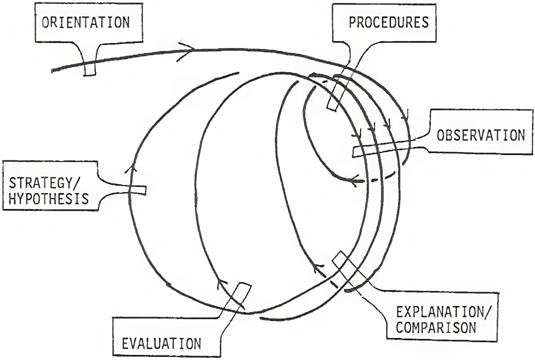
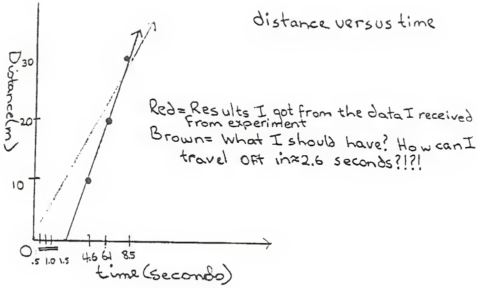


FIGURE 6-4. Progression of students' cognitive engagement with MBL activities throughout the class period.

Time for more examples

In addition to spending more time on interpretation of each event than traditional (non-MBL) laboratories allow, MBL students also had time to monitor multiple motion events. The MBL unit encouraged reiteration and replication, and on-the-spot examination of the data. The opportunity to do so is provided by the time freed from procedural activities. Students used this opportunity in several different ways.



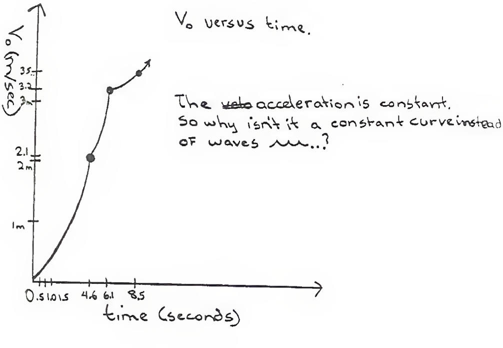


FIGURE 6-5. Graphs of data (velocity and distance) obtained from a tenth-grade student report of an inclined-plane experiment using non-MBL techniques.

They could repeat a specific event, either to verify the data, or to improve it (remove anomalies and irregularities). The practice afforded by repetition seemed to strengthen their cognitive linkage between the specific event and the graph, and also to strengthen their confidence in understanding this linkage.

They could make simple adjustments to a basic event (e.g., going a little faster, starting a bit sooner). From this experience, they would generalize from a small number of specific events to a general category of events. The repeated association of similar graphs with similar events seemed to provide students with a sense of fundamental similarities (e.g., walking away from the detector makes a distance graph with positive slope). The reinforcement could build a kind of cognitive template for representing the general class of events.

As students monitored different (new) events, they were extending their database, applying their understanding to a new situation, or testing their understanding. In effect, they were often isolating variables and identifying how they affected the system under study and the graphic representation of it. They seemed to be attending to distinctions and fundamental differences between general classes of events. The behavior of the students as they went through successive trials, isolating variables or attributes, and identifying how they were similar or different for changing events, fulfilled the major steps for inductive concept attainment described by Bruner (Joyce & Weil, 1980).

In many traditional laboratory activities, there is simply not enough time to repeat experiments. Students often have only one chance to collect and analyze their data. If something goes wrong, they

frequently copy the results from someone else. In some cases, students may not know that there was something wrong with the data until long after the experiment is over, if at all.

Making a hard task easier

In the final series of activities, the MBL students tried to reproduce a given graph composed from a sequence of five simple movements. This task was by no means trivial. It really challenged students' understanding of distance and velocity, addressing misconceptions and difficulties that plague even graduate students who teach physical science. Even when the students understood the concepts and the appropriate strategies for generating the graph, it was quite difficult for them to execute because they had to control distance and time simultaneously.

When initially faced with this difficult, complex task, students frequently claimed that they couldn't do it or didn't know how. Because they had the time to repeat experiments, they learned to break the complex event down into conceptual tasks that were more easily observed, thought about, explained and understood. By attending to these simpler tasks sequentially, they found that indeed they could do the task.

MBL seemed to encourage students to generate different routes for achieving a given goal. They used four main ways of approaching a complex task through repeated trials. Probably the most common was to approach the goal by successive improvement (shaping) through fairly minor alterations from one trial to the next. One group of tenth grade

students in the pilot study repeated this task 27 times before they were satisfied! Other students made major strategical alterations in successive trials, generally based on their conceptual understanding, but occasionally in a more haphazard manner. Other students separated the complex task into components, performed them separately, and then chained them together. Alternatively, students could concentrate on one variable at a time. For instance, they sometimes tried to reproduce the shape of the graph first (concept), and then pay attention to the scales of first one axis and then both (isolating variables). These idiosyncratic methods of simplifying a complex task provided students with a means of individualizing the conceptual chunks to the size and type that was appropriate to their understanding and aptitudes, and thus providing multiple difficulty levels and multiple ways of achieving the goal (Lepper, 1985; Malone, 1981).

Making the abstract concrete

In this MBL unit, the students were dealing with graphs (an abstract representational system) of velocity (an abstract physical property). The real-time graphing feature made the graphs responsive, manipulable, and reproducible. It seemed to make the abstract and theoretical properties of the system under study behave as though they were concrete. It has been suggested that, because graphing skills generally depend on abstract reasoning ability (McKenzie & Padilla, 1984), this ability of MBL to make abstract properties behave as though they were concrete provides students with a bridge linking the abstract

property with a graph of it (Mokros, 1986; Mokros & Tinker, in press; Thornton, 1986).

The quality and quantity of data

Data are not just easier to obtain with MBL, they are also easier to interpret because they are generally better in both quantity and quality than data obtained from traditional (non-MBL) laboratories. The MBL activity ensured that the experimental data were plentiful and of high quality, it allowed for more replication, and it enabled detailed analysis of data collected over a very short time interval.

In a traditional kinematics laboratory activity, tenth-grade students estimated distance, velocity, and acceleration as one student coasted down a hill, accelerating as a result of gravitational force. The students obtained three data points for each event (Fig. 6-5). Considering the experimental errors involved, this is inadequate for the students to use as evidence that the velocity graph is linear and goes through zero velocity, and that the distance graph is curvilinear, which surely are among the goals of the experiment. Compare the interpretability of data obtained from a high ability student in this traditional laboratory activity (Fig. 6-5) with the data obtained for a similar inclined-plane event by an MBL motion detector (Fig. 6-6). Improving the quality and quantity of data reduces the dysfunctional confusion and uncertainty of interpretation. The cognitive burden is moved from the semantic properties of the data to the syntactic

properties. Unproductive cognitive processing is removed without removing the experimental variability of the data.

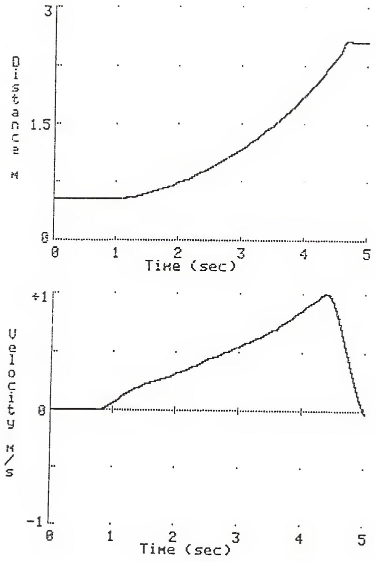


FIGURE 6-6. Graphs of data (velocity and distance) obtained from an inclined-plane experiment using MBL.

Motivational Indicators--Standard MBL

A sense of purpose

In many non-MBL activities, time constraints allow only one experiment for a given concept or concept area. In that situation the processes of evaluation and interpretation function as the appendix to the laboratory activity, rather than as the starting point for the next experiment. Their purpose is not necessarily self-evident. Evaluation is often transformed into estimation of errors--just another calculation to be performed.

In the MBL activities, the critical processes of interpretation and evaluation of results had a definite function. Students could see that these processes were integral to the progression of the activity. The data analyses and interpretations provided the bases for future action and decisions and the rationale for subsequent experiments.

It was interesting to see the changes in students' criteria for evaluating whether or not their results and their explanations were adequate. Initially, the automatic response of students was "That's OK. What's next?" As the period progressed, there was an increasing frequency of critical evaluations, self-corrections, attempts at clarification, and requests for explanations from other group members. These behaviors demonstrated that most of the students in this MBL had understood and internalized the purposes of the activity. Students seemed to learn quickly that attention to these parts of the activity

had a purpose, that they didn't cost much effort, and that they certainly made later activities easier.

A sense of autonomy

In MBL activities, students controlled the event or experiment being monitored, the computer which was doing the monitoring, and the format for representing the data. The computer was the "flunky," doing doing all the procedural tasks--measurements, calculations, and manipulations. This altered the functional roles of students compared with their roles in conventional laboratories. The delegation of tedious tasks to the computer seemed to make students feel autonomous, in control of the activity, and willing to become engaged. This contrasts with more traditional labs where students perceive the teacher to be in control of their activities, and themselves to be some kind of lackey or laborer. Students' perceptions of control and self-determination are considered to be important components of intrinsic motivation (Lepper, 1985).

A sense of competence

The MBL activity seemed to contribute to a sense of individual competence among students. Most of the students thought that the single class period with a motion detector activity had improved their understanding of both graphs and the concepts of distance and velocity. It was not uncommon to hear students, particularly the tenth graders in

the pilot study, leaving an MBL activity say with real pride of accomplishment, "That was really difficult!" This overt display of accomplishment and mastery is a strong indicator of positive motivation (Harter, 1978, 1981; Lepper, 1985; White, 1959).

A sense of curiosity

In addition to the behaviors previously mentioned (explaining, interpreting, and evaluating their results), there were a number of behavioral indicators of a strongly positive motivation. Students showed persistence (continuing the activity into their lunch hour), reduced reliance on correct procedures and worksheets, curiosity in generating hypotheses ("if ..., then ..."), and experimenting ("what if ..."). Furthermore, it was not just the frequency of these behaviors that improved. The quality of these behaviors was even more encouraging. There were noticeable improvements within a single class period in the explicitness and rigor of criteria for evaluation and in the quality of explanations and elaborations that occurred among groups of students.

Behavioral Differences with Delayed-MBL Students

The delay in graph display for the delayed-MBL students was generally 20-30 seconds for each trial run. This was not significant enough to have caused a serious difference between treatments in the opportunities for practice, repetition, or interpretation.

Nevertheless, there were noticeable differences in the behavior of the students. The differences were noticeable not because they were so large, but simply because they were so consistent in nearly every classroom.

Students in the delayed-MBL treatment seemed to be less active in their cognitive engagement with the tasks. Their conversations indicated that they continued to pay considerable attention to procedural issues throughout the class period, rather than moving their focus to interpretation and related activities. Evaluative comments were less prevalent, less rigorous, and seldom integrated into explanations. The students seemed to generate and test fewer explicit hypotheses. In other words, their activities generally did not go beyond what the worksheets requested. They did not appear to have a clear concept of the purpose and direction of the activities.

These students tended to make repetitions more because they could not remember what had happened than as attempts to improve the graph or clarify their conceptual understanding. It was not clear why they did not consider the repetitions worth their while--whether they perceived the effort to be too great or the rewards too few. It is likely to have been related to the comparative lack of manipulability and responsiveness of the system compared with the standard-MBL treatment.

There were no indications that students used the delay for any kind of maintenance or elaboration rehearsal. One such strategy is verbal mediation. Regression analysis showed that posttest performance was not significantly related to verbal ability of the students. Students seemed to treat the delay as a time to do nothing but wait. There were

no suggestions that their cognitive capacity was full to overflowing, so this attitude could have been because they either did not know or did not care to rehearse the information. The students generally seemed to have poor observational skills, relying heavily on the displayed results without referring them to the event. Indeed this seemed to be a major obstacle in generating the link between event and graph. It would be interesting to see how students would behave during a similar delayed-graph activity where they had previously received instruction in appropriate mathemagenic techniques--observation and rehearsal. I would anticipate considerable improvements in behavior, attitude, and posttest performance.

Another behavioral difference between the two groups was that the students in the delayed-MBL groups were more like receptors of information than investigators. They seemed to use the system to obtain data, but seldom progressed to using the system to ask questions and seek information to guide explanations and performance. To use the conceptual framework for graph comprehension discussed in the literature review, these students recognized the conceptual message of the displayed data (bottom-up encoding), but seldom imposed conceptual questions on the experimental situation (top-down encoding).

The difference in behavior between the students in the two MBL groups is clearly described by Wertime's (1979) analysis of the attitudinal components of thinking and problem-solving, although again it should be noted that the differences were consistent rather than large. The standard-MBL students are identifed by the state he describes where the self and the task become an inseparable whole, with

students hopeful, forward looking, and attention outwards towards the task. The task appears light and flexible to the end of the class period. In contrast, the students in the delayed-MBL treatment never really became inseparable from the task, and the task was not fully internalized. The difference was most noticeable towards the end of the class period, as the standard-MBL students seemed to be getting gradually more absorbed with the activity, whereas the delayed-MBL students exhibited progressively greater detachment from the task.

In the delayed-MBL treatment groups, there seemed to be fewer indications of humor, fun, and playfulness with the equipment. There was less conversation of any kind, and what conversation there was seemed to be more tentative and less assured. These students were less persistent, sometimes disengaging from the activities before they were instructed to do so. There were instances when these students made comments like "What are they (standard-MBL group) laughing at?", and students in a couple of groups leaned on the bench watching the other group. (They could not see the other group's computer screen.) They certainly seemed to feel that the "grass was greener" with the standard-MBL group. In contrast, several groups of standard-MBL students had to be told repeatedly to start packing up, and one group stayed into their lunch break in order to complete the supplementary questions on the worksheet (i.e., not even to work with the MBL activity).

A further indication of the differences in level of engagement of students in the different treatments can be seen in the comments volunteered on the posttreatment questionnaire (Appendix G). There was a decrease from the standard-MBL to the delayed-MBL to the pencil-and-

paper students in the number of comments, the length of the comments, and their relevance to their treatment activity rather than to the research.

Summary

The automation of routine procedures allowed students to spend their time and attention on higher order cognitive and perceptual tasks. This appeared to contribute to the development of critical thinking, problem-solving skills, and an understanding of science processes. The real-time feature seemed to have a considerable impact on both cognitive and motivational behavior, encouraging students to feel more autonomous, more competent, and more free to experiment. This seems to have been more pronounced for males, who rated the standard-MBL activity more interesting and less difficult, than did females in the same treatment or both sexes in other treatments.

CHAPTER VII

CONCLUSION

Summary of Dissertation Research

Microcomputer-based laboratory (MBL) activities, as used in this study, involve the use of the computer coupled with an appropriate sensor to automate the procedures of collecting data, performing calculations, and displaying the results in graph format. Data are collected by a motion-detector sensor and may be displayed as either a distance-time graph or a velocity-time graph.

Research Objectives

This research examined the effectiveness of a specific MBL activity in improving students' conceptual understanding and graphing skills. It focused on both MBL as tool and students' ability to interpret kinematics graphs as outcome. The reseach was designed so that, once having established that performance does improve after an MBL activity, the results could be examined for evidence of (a) what kind of learning took place (graphing or kinematics concepts), (b) what kind of problems did students have with graphing and with the concepts, (c) what kinds of students had problems (ability, gender, etc.), and (d) what specific

characteristics of the MBL tool facilitated the learning (focusing specifically on real-time data display).

Other studies (Linn et al., in press; Mokros & Tinker, in press) have demonstrated improvements with performance in interpreting graphs at the middle-school level following extended experience with MBL activities. This research, with a single class period of MBL activity, was intended to complement such studies by exploring the other end of the treatment-time spectrum, by using older students, and by examining in finer grain the mechanism of how and why MBL helps students to develop the cognitive linkage between a physical event and a graph.

Research Design

The research concentrated on real-time graphing as the most salient feature of the MBL tool. The impact of the immediate graph display of data in standard MBL was isolated by delaying the graph display seen by students until after the data had been collected, a delay of about 20-30 seconds, but leaving the activity the otherwise identical. These two experimental treatments (standard and delayed MBL) were compared with a pencil-and-paper graphing activity to provide a baseline that approximated normal classroom procedures and a pretest/posttest only statistical control.

High school physics students from seven rural schools participated in the study. After a day of pretesting, students were randomly assigned within classes to the four treatments described above for a single class period of treatment. They worked on the kinematics

activity in small groups of two or three students in the MBL treatments and two to four students for the pencil-and-paper activity. Posttreatment testing took place the following day.

This research used the usual battery of covariates, attributes of the learner, used to explain variance in many kinds of learning and performance (i.e., sex, age, general ability, reasoning and development). Graphing skills are difficult to measure and the literature does not provide an adequate instrument for assessing graphing skills and revealing the underlying concepts that students have about graphs and graphing (i.e., their graph schema). One reason for this seems to be the difficulty of separating the conceptual understanding of graphs, as a system of representation, from that of the concepts being represented.

In addition to a content-specific posttest interpreting graphs of motion, students completed tests of ability to interpret graphs of a wide range of physical phenomena and a test of ability to construct a graph from experimental data provided for them. Results from these tests were examined with the intention of providing guidance for explaining experimental results and for designing future research. It was not expected that a single class experience with MBL would produce significant differences on a general graphing measure but it might generate improvement with content-specific graphs.

Quantitative Research Results

The posttest required students to translate between a verbal description and a graphic representation of given motion events. A factorial analysis of covariance determined that students in the standard-MBL treatment group had significantly ($p<.05$) lower error rates on the posttest than did students in either the delayed-MBL or the pencil-and-paper treatments. When the posttest was subdivided, the same pattern was found with distance-time items, but there was no statistically significant difference on velocity-time items.

Most of the posttest difference appears to be attributable to the real-time graphing feature of the standard-MBL treatment. This interpretation should be received with some caution because students in the delayed-MBL treatment had lower general ability (SAT scores) and included a disproportionately high number of females.

In general, females were younger than males, and, even allowing for this difference, had lower scores on SAT and development. After taking all these factors into consideration, females had greater difficulty with graphs of velocity, but not with graphs of distance. They reported on questionnaires that they had more difficulty and less enjoyment with both graphs and the MBL activity than did males. After the standard-MBL treatment, females made substantial improvement in translating graphs of distance, whereas males, who had generally reached a performance ceiling with distance graphs, increased their advantage with velocity graphs.

Qualitative Research Results

Observations of the students' behavior as they participated in the treatment activities suggest that there were substantial differences in the quality of cognitive engagement of the two MBL treatment groups. Because there was very little difference between the opportunities provided in each treatment, it seems that the behavioral differences arose predominantly from motivational differences. Alternative explanations are that the students in the delayed-MBL treatment group may have lacked appropriate mathemagenic techniques to profit by the delayed graph display, or the additional cognitive burden imposed by the delay may have exceeded their cognitive capacity and been dysfunctional.

There was time and opportunity for interpretation and confirming experiments. Throughout the class period, in the standard-MBL groups the standards for evaluating data became more explicit and more rigorous, the explanations became more comprehensive, and the language exhibited improvements in conceptual and grammatical clarity. The most exciting potential for MBL appeared to be its impact on critical thinking and communication skills, and on competence motivation.

Questionnaire ratings also indicated that all students found the activities intrinsically interesting, even those who had previously indicated a lack of interest in graphs per se. The standard-MBL activity appeared to promote positive social interaction as shown by higher ratings of group members of being helpful in the activity.

Interpretation of Results

Students in the standard-MBL treatment appeared to improve their understanding of the conventions of graphing and of the concept of distance. There was no evidence of improvement with the concept of velocity.

Although the students generally thought that they had little difficulty with either interpreting graphs or with constructing graphs, their performance on a number of measures indicated a superficial competence with graphing. In general, students had little difficulty interpreting graphs of simple properties, but errors increased dramatically as the content of the graph became conceptually more complex, more abstract, or conflicted with the visual image generated by the description of the graph.

A test of constructing a graph from raw data (where time was not a relevant variable) revealed that about one fifth of the students had serious functional inadequacies with their graph schemas, not knowing how to represent information on a graph to demonstrate its underlying pattern. Among these students, females were comparatively low on measures of ability, but there was no clear explanation for these males being unable to construct an adequate graph. The frequency of nonfunctional graph schemas may well explain why students fail to learn graphing skills from inductive MBL experiences, which do not explicitly address the relevant graph schema.

Future Research

Tentative as the results must be from research with such a short treatment period, they nevertheless provide suggestions of several issues that would be fruitful for future research. These consist of questions not addressed by this study and questions not answered. They also include tentative interpretations that need to be examined. One such example is the need for a less ambiguous distinction between evidence of learning concepts and learning graphing conventions.

Having identified real-time graphing as an important, and probably the main, feature of MBL that facilitates the immediate cognitive linkage between event and graph, further research grounded in cognitive psychology could explore the information-processing mechanisms involved. It would be particularly interesting to find out why some students persist in making basic errors after extended MBL experience (Linn et al., in press; Mokros & Tinker, in press).

Two issues evolved from the analysis of what kinds of problems students have, and what kind of students have problems. This research identified sex as an important factor in graphing skills, but failed to indicate whether MBL experience would selectively improve performance on a gender basis. A longer treatment period and a range of content would be needed to address this question. It is also uncertain to what extent graphing skills rely on spatial visualization skills, which generally favor males.

The second issue arises from the apparent prevalence of students with inappropriate or inadequate graph schemas. This problem may

underly the core of students who continued to have fundamental problems in interpreting graphs after extended MBL experience. Future research needs to characterize such problems, to explore their effect on learning inductively from MBL experience, and to determine how they can be remediated.

As part of the characterization process, it would be productive to use eye fixation techniques to compare the way people experienced in using graphs extract information from them, using content with which they are either familiar or not. There would probably be substantial differences in the sequence and duration of attention to the referents of the graph (label, axes, scales) compared with the data.

At least part of the difficulty students had with comprehension of kinematics graphs appeared to be an inability to isolate and focus on the relevant concepts. They did not seem to be aware of the importance of identifying the dependent variable and relating it to the independent variable. This is one area of basic skills development where explicit instruction in mathemagenic techniques may be both appropriate and effective.

Another such area is in the use of memory maintenance and rehearsal techniques during the delay between the occurrence of an event and the manifestation of it. Developing such skills should improve the clarity of students' observations of experimental events, their expectations of outcomes, and their formulation of hypotheses.

Patterns of data in this study raise some interesting possibilities that could fruitfully be explored in future research. One is the use of resilience of errors as an indication of conceptual stability, and as it

correlates with mathemagenic behavior and with attributes of the individual student. Another is the use of performance in bidirectional translations as an index of automaticity of cognitive linking of concepts.

Implementation Considerations

MBL makes available to the high school techniques common in industry and research today. It allows us to measure physical properties that were otherwise difficult, if not impractical, to measure. It also allows for data to be collected during intervals of time that can be either very long or very short (split seconds) simply by selecting different data-sampling intervals, although this factor was not investigated in this study.

Throughout the activity, the student remains in control of such experimental decisions as the specific event being monitored, the parameter displayed, and some of the characteristics of the graph display. Various MBL units allow students to represent their data as histograms or Cartesian graphs, to alter the scales on the axes, to examine each data point with the option of excluding it from subsequent data analysis, and to compare two sets of data (graphs). These options provide students with unprecedented ability to examine how changes in the representation of data help or hinder the process of interpretation. This experience with the general representational properties or attributes of the graphs should contribute to their development of an

appropriate "graph schema"--an understanding of what graphs are for, what they do, and how they store and display information.

The automation of routine procedures allows students to spend their time and attention on higher order cognitive and perceptual tasks. These tasks have long been regarded by researchers as important in developing critical thinking, problem-solving skills, and in developing an understanding of the science processes. Practical considerations have restricted the ability of teachers to incorporate them into the classroom. As MBLs address this different set of objectives, they seem to reveal poor science-process skills, habits of passivity, and major deficiencies in basic skills such as graphing and critical thinking.

Students are known to have particular difficulty in distinguishing between distance and velocity because of their intuitive conceptions developed from prior experiences in the course of living. These misconceptions are problematic because they are so very stable and resistant to change. In early stages of the treatment activities, it was common for students to attend to the wrong part of a graph, and reinforce their misconception. For instance, many of them expected a velocity-time graph of a constant velocity event to have a positive slope. As the "mover" started to move from standstill (i.e., accelerated), the slope of the graph produced was positive--until the mover reached steady velocity. The students confused this acceleration period with constant velocity.

As is the case with any educational technique, MBL can help or hinder the development of concepts. It was possible for students to misuse MBL to verify, or at least accomodate, their misconceptions.

Alternatively, they could change their conceptual understanding. This self-correction generally occurred spontaneously as students progressed through the final phase of the activity. Teachers need to be aware of such potential trouble spots and not allow students to skip over them. Attention to such trouble spots needs to be specific, explicit, and repeated.

One of the effective features of laboratory units developed by TERC is that learning occurs by the parallel development of the concepts of the physical property being studied and the graph. As an example, students' initial understanding of graphs is employed in learning the properties of velocity and distance. At the same time, students' understanding of speed and distance is used in learning the properties of the graph. It is uncertain how effective this inductive approach would be for those students who have both an inadequate understanding of the function of a graph (a nonfunctional graph schema), and substantial misconceptions about the physical property being studied. In this study, there were insufficient of these students to address this question.

Some science people argue that MBL automates things that should not be automated, particularly with regard to graphing and calculating--the "black box" argument. There has been some debate about who should do the graphing--computer or student. In general, research indicates that, by the time they get to high school, most students can plot data on a graph and can read data points from a graph, although they are less capable of assigning variables to the axes or of selecting appropriate scales (discussed in Chapter II). The graphing skill most consistently

lacking is the ability to interpret graphs, and this is where MBLs appear to provide the most practice. Evidence from both the posttest performance and the behavioral observations indicates that students were attending to and learning the properties of graphs--scales, slopes, and relationships. The "black box" fear seems to be more relevant to laboratory simulations than to MBLs.

MBL activities are not able to, nor intended to, replace conventional laboratories entirely, nor is it intended to eliminate completely the students' need to perform calculations and construct graphs within science courses. They are designed to be a valuable addition to the repertoire of ways to teach and learn science.

Conclusion

MBL activities have the potential to make a beneficial contribution to curriculum and instruction in science education. They allow students to spend more time and attention on higher order cognitive and perceptual tasks, and less time on procedural and manual tasks. They seem to contribute simultaneously to the development of science concepts, graphing skills, and the ability to use (rather than merely understand) high level science processes. Results from this study indicate that most of the facilitative effects of MBLs reside in the immediacy of display of the real-time graphing feature.

APPENDIX A

WORKSHEET--MBL TREATMENTS

<u>Distance graphing</u>

Take turns using the computer to make the following DISTANCE graphs:
a) walk slowly away from the detector, then slowly towards it.
b) walk medium speed towards the detector, then medium speed away.

Describe (in words or sketch) what happens to the slope of the graph as you change
a) velocity ?
zero (stand still)

slow

faster

b) direction?
towards

turn around

away

<u>Velocity graphing</u>

Take turns using the computer to make the following VELOCITY graphs:
a) walk slowly away from the detector, then slowly towards it.
b) walk medium speed towards the detector, then medium speed away.

Are the slopes of the constant velocity flat or sloping?

Describe (in words or sketch) what happens to the slope of the graph as you change
a) velocity ?
zero (stand still)

slow

faster

b) direction?
towards

turn around

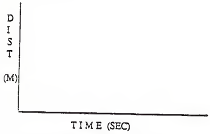
away

Predicting Graphs

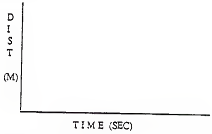
Draw below your prediction of the DISTANCE graph of a person
a) starts at 1 meter,
b) walks steadily away covering 3 meters in 3 seconds,
c) stops for 3 seconds, and
d) walks slowly back covering 3 meters in 6 seconds.

Compare predictions. Make sure you all agree, then do the experiment until you get a good graph. Sketch it, including time and distance scales.

PREDICTION



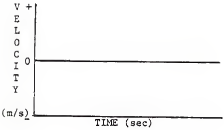
EXPERIMENT



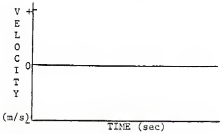
Draw your prediction of the VELOCITY graph of the same event.

Compare predictions. Make sure you all agree, then do the experiment until you get a good graph. Sketch it, including time and distance scales.

PREDICTION



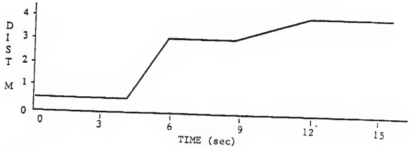
EXPERIMENT

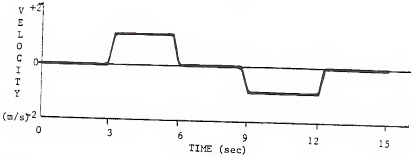


Questions:
1. What was the average velocity at b?
2. What was the average velocity at c?
3. What was the average velocity at d?
4. What is the relationship between slope of the distance graph and velocity?
Remember: velocity is the rate of change of distance over time.

Creating Graphs

a) Your challenge is to try to reproduce the graphs below.
b) Try hard to get nice flat sections and the times right.
c) Everyone in the group should make a good copy of the graphs.
d) Show on the graphs what you did to make them.
e) Have the experimenter check your graphs when you finish.





Questions:

1. On the velocity graph, can you tell where an object started?

2. When you are walking at constant velocity--
   Are distance graphs flat or sloping?
   Are velocity graphs flat or sloping?

3. Why can you never get a vertical line on a velocity graph?

4. What is the instantaneous velocity (in m/sec) when you turn around?

5. How is acceleration (rate of change of velocity over time) represented on a velocity graph?

APPENDIX B

WORKSHEET--PENCIL AND PAPER TREATMENT

Drawing graphs

Imagine you walk at a constant velocity of 1 m/s away from a point for four seconds, then stop for four seconds, and walk back to the start at a constant velocity of 2 m/s.

(1) Draw a graph of your distance from the starting point. Use the graph paper provided.

(2) Draw a graph of your velocity. [Remember: velocity is not the same as speed. It can be positive or negative.]

Imagine you are four meters from your desk. You stand there for three seconds, then walk towards the desk at a steady velocity of 1 m/s for three seconds, stand there for two seconds, then walk away at a steady velocity of 2 m/s for two seconds.

(3) Draw a graph of your distance from the desk.

(4) Draw a graph of your velocity. [Use the convention that velocity is positive when distance from the desk is increasing.

Compare graphs. Do you all agree?

Questions:

Discuss the following questions:

(1) Compare your distance and velocity graphs.
When you are walking at constant velocity--
Are the distance graphs flat or sloping?
Are the velocity graphs flat or sloping?

(2) What happens to the slope of your distance graph when you go faster? slower? stop?

(3) What happens to the slope of your velocity graph when you go faster? slower? stop?

(4) Can you tell from a velocity graph where you started from?

(5) If you walk away from a point, turn around and come back, what is your velocity when you turn around?

Interpreting graphs

A man is driving home from work. The graph below shows his speed (not velocity) during the entire trip.

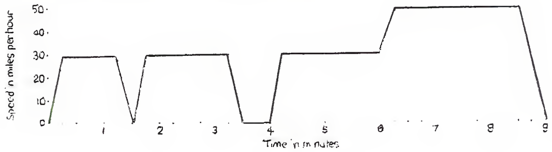
Make up a story to explain the graph. In your story tell--

a) Approximately how far he lives from where he works,

b) What happened one and one-half minutes after he started,

c) What happened between three and one-half and four minutes after he started, and

d) Something observed about speed limits.



APPENDIX C

PRETEST--DISTANCE AND VELOCITY GRAPHS

DISTANCE GRAPHING

Choose the correct distance/time graph for each of the following questions. You may use a graph more than once or not at all.

1. Which distance graph represents the motion of an object with a steady velocity away from the detector?

   ____

2. Which distance graph is one of an object standing still?

   ____

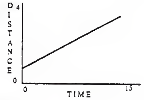
3. Which distance graph is one of an object with a steady speed toward the detector?

   ____

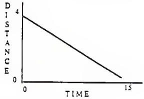
4. Which distance graph shows that the object turned around?

   ____

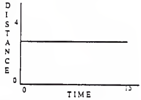
A



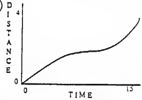
B



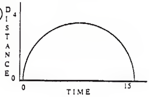
C



D



E



VELOCITY GRAPHS

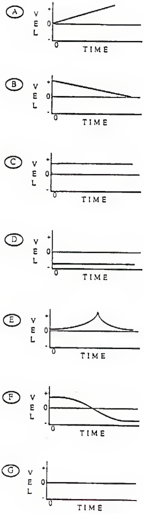
4. Which velocity graph shows an object going away from the detector at a steady velocity? ________

5. Which velocity graph shows an object that is standing still? ________

6. Which velocity graph shows an object moving toward the detector at a steady velocity? ________

7. Which velocity graph shows an object changing direction? ________

8. Which velocity graph shows an object that is steadily increasing its speed (accelerating)? ________



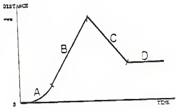
APPENDIX D

POSTTEST--DISTANCE AND VELOCITY GRAPHS

In the study, this posttest was presented in a booklet, with each question presented on a separate page.

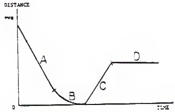
Which section of this DISTANCE graph represents an object 1

with constant positive velocity?



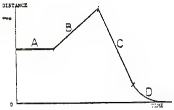
Which section of this DISTANCE graph represents an object 2

at rest?



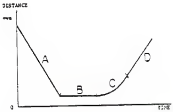
What section of this DISTANCE graph represents an object 3

moving steadily towards a detector?



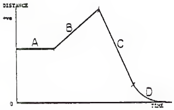
Which section of this DISTANCE graph represents an object 4

moving steadily away from a detector?

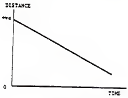


Which section of this DISTANCE graph represents an object 5

with constant negative velocity?



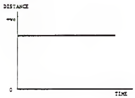
6



This DISTANCE graph represents an object

A. at rest
B moving steadily away from a detector
C. moving steadily towards a detector
D. turning around

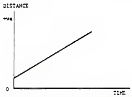
7



This DISTANCE graph represents an object

A. moving steadily away from a detector
B. at rest
C. turning around
D. moving steadily towards a detector

8

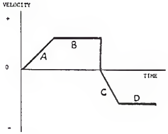


This DISTANCE graph represents an object

A. turning around
B. moving steadily towards a detector
C. moving steadily away from a detector
D. at rest

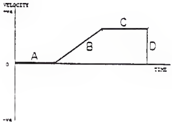
Which section of this VELOCITY graph represents an object 9

with constant negative velocity?



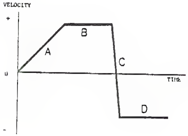
Which section of this VELOCITY graph represents an object 10

accelerating (increasing velocity)?



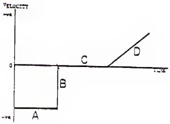
Which section of this velocity graph represents an object 11

moving steadily away from a detector?



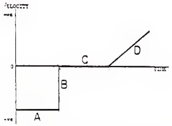
Which section of this VELOCITY graph represents an object 12

stopping almost instantaneously?



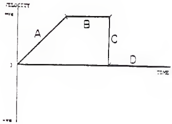
13 Which section of this VELOCITY graph represents an object

at rest?



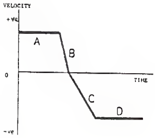
14 Which section of this VELOCITY graph represents an object

with constant positive velocity?



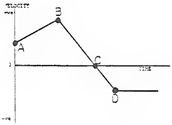
15 Which section of this VELOCITY graph represents an object

moving steadily towards a detector?



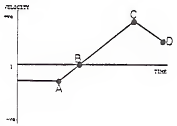
16 Which point on this VELOCITY graph represents an object

with negative velocity?



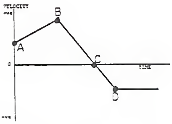
17 Which point on this VELOCITY graph represents an object

turning around?



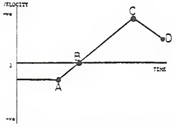
18 Which point on this VELOCITY graph represents an object

with zero velocity?

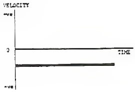


19 Which point on this VELOCITY graph represents an object

with maximum velocity?



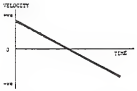
20



This VELOCITY graph represents an object

A. with constant positive velocity
B. with constant negative velocity
C. at rest
D. accelerating (increasing velocity)

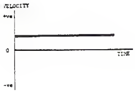
21



This VELOCITY graph represents an object

A. moving steadily away from a detector
B. turning around
C. moving steadily towards a detector
D. with constant negative velocity

22



This VELOCITY graph represents an object

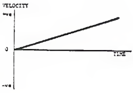
A. turning around
B. moving steadily towards a detector
C. at rest
D. moving steadily away from a detector

23

This VELOCITY graph represents an object

A. at rest
B. moving steadily away from a detector
C. moving steadily towards a detector
D. accelerating (increasing velocity)

24



This VELOCITY graph represents an object

A. with constant positive velocity
B. moving steadily towards a detector
C. turning around
D. accelerating (increasing velocity)

APPENDIX E

QUESTIONNAIRE--GRAPHS IN SCIENCE COURSES

Instructions:
For each of the following questions, please circle the word that best represents your answer. There are no "correct" or "incorrect" answers.

1. Are graphs useful in your science courses?

   Never Not often Sometimes Mostly Always

2. Do graphs make science more interesting?

   Never Not often Sometimes Mostly Always

3. Do you have difficulty understanding graphs?

   Never Not often Sometimes Mostly Always

4. Do you have difficulty constructing graphs?

   Never Not often Sometimes Mostly Always

5. Do you enjoy working with graphs?

   Never Not often Sometimes Mostly Always

6. How often do you use graphs voluntarily in explaining things in science?

   Never Not often Sometimes Mostly Always

APPENDIX F

QUESTIONNAIRE--EXPERIMENTAL ACTIVITY

Instructions: For each of the following questions, please circle the response that best represents your answer. There are no "correct" or "incorrect" answers.

1. Was this activity <u>useful</u>?

   Not at all   No   Undecided   Yes   Very much

2. Was this activity <u>difficult</u>?

   Not at all   No   Undecided   Yes   Very much

3. Was this activity <u>interesting</u>?

   Not at all   No   Undecided   Yes   Very much

4. Was this activity <u>confusing</u>?

   Not at all   No   Undecided   Yes   Very much

5. Did you <u>enjoy</u> this activity?

   Not at all   No   Undecided   Yes   Very much

6. Did this activity help you understand <u>motion</u>?

   Not at all   No   Undecided   Yes   Very much

7. Did this activity help you understand <u>graphs of motion</u>?

   Not at all   No   Undecided   Yes   Very much

8. How <u>helpful</u> were other members in your group?

   Not at all   No   Undecided   Yes   Very much

9. Were there other things that you would like to have done in this activity?

   Not at all   No   Undecided   Yes   Very much

APPENDIX G

STUDENTS' COMMENTS ON POSTTREATMENT QUESTIONNAIRE

Standard MBL (comments from 9 students)

1 (Male) It was a good idea for teaching methods, but there was still something missing.

26 (Female) I wish I could've finished the computer part of it. I think I would have understood a lot more. It was fun, too.

34 (Male) It was interesting to see immediate graph results as we analyzed the motion of an object. Would be interesting to see the graph as the "sonic" detector was trained on an object undergoing circular motion.

36 (Male) The activity was a good one. It showed the relationship of time to distance and velocity graphs. It helped me understand graphs better.

52 (Female) It made graphs easier to understand and it was also fun.

56 (Female) I thought this activity was fun an interesting. It helped me understand graphs more. Although I don't really understand them as well as I would like to.

65 (Male) Learned how the equipment worked.

82 (Female) Very well organized.

87 (Female) It was interesting.

Delayed MBL (comments from 7 students)

15 (Female) Your experiment was fun and exciting learning about motion and I hope that you do good on your PhD.

21 (Female) I think if we had more time it would be easier to learn by this method.

30 (Female) This was an interesting three days and I enjoyed it.

33 (Female) It was a learning experience.

47 (Male) It was a very good course that helped me to <u>like</u> and <u>understand</u> graphs better.

77 (Male) This activity challenged my mind to remember what I had learned about distance and velocity graphs.

80 (Male) Interesting and challenging.

<u>Pencil and Paper</u> (comments from 6 students)

14 (Male) It was interesting to see something that I thought I knew a lot about and to find I knew very little.

29 (Male) I basically enjoyed this. There were some parts that were confusing.

37 (Female) This activity was interesting and fun. I enjoyed answering the questions about graphs on the computers.

39 (Female) Thanks for saving us from three days of his lectures!

54 (Female) <u>Interesting!</u>

71 (Male) More use on the computers should have been done by the students.

<u>Test Only</u> (comments from 3 students)

5 (Male) Wish I had more time to do the middle part of this test.

67 (Male) Well planned, well set-up experiments.

81 (Female) I wish I would have been here the first day so I could have been informed more on the activity.

REFERENCES

Anscombe, F. J. (1973). Graphs in statistical analysis. American Statistician, 27, 17-21.

Arons, A. B. (1979). Some thoughts on reasoning capacities implicitly expected of college students. In J. Lochhead & J. Clement (Eds.), Cognitive process instruction: Research on teaching thinking skills (pp. 209-215). Philadelphia: Franklin Institute Press.

Ausubel, D. P. (1968). Educational Psychology. New York: Holt, Rhinehart, & Winston.

Baker, S. R., & Talley, L. H. (1972). The relationship of visualization skills to achievement in freshman chemistry. Journal of Chemical Education, 49, 775-776.

Baker, S. R., & Talley, L. H. (1974). Visualization skills as a component of aptitude for chemistry--A construct validation study. Journal of Research in Science Teaching, 11, 95-97.

Barclay, W. L., III. (1986, June). Graphing misconceptions and possible remedies using microcomputer-based labs. Paper presented at the 7th National Educational Computing Conference, University of San Diego, San Diego, CA.

Berlyne, D. E. (1957). Conflict and information-theory variables as determinants of human perceptual curiosity. Journal of Experimental Psychology, 53, 399-405.

Berlyne, D. E. (1966). Curiosity and exploration. Science, 153, 25-33.

Berlyne, D. E. (1978). Curiosity and learning. Motivation and Emotion, 2(2), 97-175.

Bertin, J. (1981). Graphics and graphic information-processing. (W. J. Berg & P. Scott, Trans.). Berlin: Walter de Gruyter & Co. (Original work published 1977)

Bertin, J. (1983). Semiology of graphics: Diagrams networks maps. (W.J. Berg, Trans.). Madison, WI: University of Wisconsin Press. (Original work published 1973)

Bestgen, B. J. (1980). Making and interpreting graphs and tables: Results and implications from national assessment. Arithmetic Teacher, 28(8), 26-29.

Bloom, B. S., & Krathwohl, D. R. (1977). Taxonomy of behavioral objectives: Handbook I: Cognitive domain. New York: Longman.

Bobrow, S. A., & Bower, G. H. (1969). Comprehension and recall of sentences. Journal of Experimental Psychology, 80, 455-461.

Bradley, R. L. (1968). Is the science laboratory necessary for general education science courses? Science Education, 52, 58-66.

Bransford, J.D. (1979). Human cognition. Learning, understanding and remembering. Belmont, CA: Wadsworth.

Bredderman, T. (1982). The effects of activity-based science in elementary schools. In M. B. Rowe (Ed.), Education in the 80's: Science (pp. 63-75). Washington, DC: National Education Association.

Bredderman, T. (1984). The influence of activity-based science programs on classroom practices: a qualitative synthesis. Journal of Research in Science Teaching, 21, 289-303.

Burstein, L. (1980). The analysis of multilevel data in educational research and evaluation. Review of Research in Education, 8, 158-233.

Chambers, J. M., Cleveland, W. S., Kleiner, B., & Tukey, P. A. (1983). Graphical methods for data analysis. Belmont, CA: Wadsworth.

Champagne, A. B., Klopfer, L. E., & Anderson, J. H. (1980). Factors influencing the learning of classical mechanics. American Journal of Physics, 48, 1074-1079.

Charen, G. (1966). Laboratory methods build attitudes. Science Education, 50, 54-57.

Ciarcia, S. (1984). An ultrasonic ranging system. Byte, 9(10), 113-123.

Clement, J. (1982). Students' preconceptions in introductory mechanics. American Journal of Physics, 50, 66-71.

Clement, J. (1985, July). Misconceptions in graphing. Paper presented at the Ninth Conference of the International Group for the Psychology of Mathematics Education, Noordwijkerhout, The Netherlands.

Clement, J., Mokros, J. R., & Schultz, K. (1986, April). Adolescents' graphing skills: A descriptive analysis. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.

Cleveland, W. S. (1984a). Graphs in scientific publications. The American Statistician, 38, 261-269.

Cleveland, W. S. (1984b). Graphical methods for data presentation: Full scale breaks, dot charts, and multibased logging. The American Statistician, 38, 270-280.

Cleveland, W. S. (1985). The elements of graphing data. Monterey, CA: Wadsworth.

Cleveland, W. S., & McGill, R. (1984). Graphical perception: Theory, experimentation, and application to the development of graphical methods. Journal of the American Statistical Association, 79, 531-554.

Cleveland, W. S., & McGill, R. (1985). Graphical perception and graphical methods for analyzing scientific data. Science, 229, 828-833.

Cohen, K. M. (1981). The development of strategies of visual search. In D. F. Fisher, R. A. Monty & J. W. Senders (Ed.), Eye movements: Cognition and visual perception (pp. 271-288). Hillsdale, NJ: Erlbaum.

Coltheart, M., Hull, E., & Slater, D. (1975). Sex differences in imagery and reading. Nature, 253, 438-440.

Condry, J. (1977). Enemies of exploration: Self-initiated versus other-initiated learning. Journal of Personality and Social Psychology, 35, 459-477.

Condry, J., & Chambers, J. (1978). Intrinsic motivation and the process of learning. In M. R. Lepper & D. Greene (Eds.), The hidden costs of reward: New perspectives on the psychology of human motivation (pp. 61-84). Hillsdale, NJ: Erlbaum.

Coulter, J.C. (1966). The effectiveness of inductive laboratory, inductive demonstration, and deductive laboratory in biology. Journal of Research in Science Teaching, 4, 185-186.

Cox, D. R. (1978). Some remarks on the role in statistics of graphical methods. Applied Statistics, 27, 4-9.

Cronbach, L. J. (with the assistance of J. E. Deken & N. Webb) (1976). Research on classrooms and schools: Formulation of questions, design and analysis (Occasional Paper). Stanford Evaluation Consortium, Stanford, CA.

Culbertson, H. M., & Powers, R. D. (1959). A study of graph comprehension difficulties. Audio-Visual Communication Review, 7, 97-110.

Deci, E. L., & Porac, J. (1978). Cognitive evaluation theory and the study of human motivation. In M. R. Lepper & D. Greene (Eds.), The hidden costs of reward: New perspectives on the psychology of human motivation (pp. 144-176). Hillsdale, NJ: Erlbaum.

Dillashaw, F. G., & Okey, J. R. (1980). Test of the integrated science process skills for secondary science students. Science Education, 64, 601-608.

diSessa, A. A. (1982). Unlearning Aristotelian physics: A study of knowledge-based learning. Cognitive Science, 6, 37-75.

Doran, R. L. (1978). Measuring the "Processes of Science" objectives. Science Education, 62, 19-30.

Dugdale, S., & Kibbey, D. (1983). Interpreting Graphs [computer program]. Conduit.

Egelston, J. (1973). Inductive vs. traditional methods of teaching high school biology laboratory experiments. Science Education, 57, 467-477.

Eggen, P., Kauchak, D., & Kirk, S. (1978). The effects of generalizations as cues on the learning of information from graphs. Journal of Educational Research, 71, 211-213.

Eliot, J., & Fralley, J. S. (1976). Sex differences in spatial abilities. Young Children, 31, 487-498.

Esler, W. K. (1982). Physiological studies of the brain: Implications for science teaching. Journal of Research in Science Teaching, 19, 795-803.

Fennema, E., & Sherman, J. (1977). Sex-related differences in mathematics achievement, spatial visualization and affective factors. American Educational Research Journal, 14, 51-71.

Fienberg, S. E. (1979). Graphical methods in statistics. The American Statistician, 33, 165-178.

Frase, L. T. (1968). Questions as aids to reading: Some research and theory. American Educational Research Journal, 5, 319-332.

Fraser, B. J. (1980). Development and validation of a test of enquiry skills. Journal of Research in Science Teaching, 17, 7-16.

French, J. W., Ekstrom, R. B., & Price, L. A. (1963). Manual for kit of reference tests for cognitive factors. Princeton, NJ: Educational Testing Service.

Fry, E. (1981). Graphical literacy. Journal of Reading, 24, 383-390.

Furth, H. (1970). An inventory of Piaget's developmental tasks. Washington, DC: Center for Research in Thinking and Language, Department of Psychology, Catholic University.

Garron, D. C. (1970). Sex-linked, recessive inheritance of spatial and numerical abilities, and Turner's syndrome. Psychological Review, 77, 147-152.

Goetz, J. P., & LeCompte, M. D. (1984). Ethnography and qualitative design in educational research. Orlando, FL: Academic Press.

Guidry, N. P., & Frye, K. B. (1968). Graphic communication in science. Washington, DC: National Science Teachers Association.

Halloun, I. A., & Hestenes, D. (1985a). The initial knowledge state of college physics students. American Journal of Physics, 53, 1043-1055.

Halloun, I. A., & Hestenes, D. (1985b). Common sense concepts about motion. American Journal of Physics, 53, 1056-1065.

Harter, S. (1978). Effectance motivation reconsidered: Toward a developmental model. Human Development, 21, 34-64.

Harter, S. (1981). A new self-report scale of intrinsic versus extrinsic orientation in the classroom: Motivational and informational components. Developmental Psychology, 17, 300-312.

Hartlage, L. C. (1970). Sex-linked inheritance of spatial ability. Perceptual and Motor Skills, 31, 610.

Hegarty, E. H. (1982). The role of laboratory work in science courses: Implications for college and high school levels. In M. B. Rowe (Ed.), Education in the 80's: Science (pp. 82-105). Washington, DC: National Education Association.

Hilton, T. L. (1985). National changes in spatial-visual ability from 1960 to 1980 (Research Report). Princeton, NJ: Educational Testing Service.

Hofstein, A., & Lunetta, V. N. (1982). The role of the laboratory in science teaching: Neglected aspects of research. Review of Educational Research, 52, 201-217.

Johnson, R. T., Ryan, F. L., & Schroeder, H. (1974). Inquiry and the development of positive attitudes. Science Education, 58, 51-56.

Jones, M. (1985, July). Getting good graphs. PC Magazine, 217-218.

Joyce, B., & Weil, M. (1980). Models of teaching (pp. 22-46). Englewood Cliffs, NJ: Prentice-Hall.

Kagan, J. (1972). Motives and development. Journal of Personality and Social Psychology, 22, 51-66.

Kamii, C. (1984, February). Autonomy: The aim of education envisioned by Piaget. Phi Delta Kappan, 65, 410-415.

Kauchak, D., Eggen, P., & Kirk, S. (1978). The effect of cue specificity on learning from graphical materials in science. Journal of Research in Science Teaching, 15, 499-503.

Kellogg, M. G. (1967). The effect of laboratory-discovery methods and demonstration-discussion methods upon elementary science methods students' abilities to analyze and interpret graphs. Dissertation Abstracts International, 27, 3345-A. (University Microfilms No. 67-4017)

Kirk, S., Kauchak, D., & Eggen, P. (1978). The effect of cue specificity and locus on learning from graphical material. Journal of Educational Research, 72, 39-44.

Kosslyn, S. M. (1985). Graphics and human information processing. Journal of American Statistical Association, 80, 499-512.

Leboutet-Barrell, L. (1976). Concepts of mechanics among young people. Physics Education, 11, 462-466.

Lefferts, R. (1981). Elements of graphics: How to prepare charts and graphs for effective reports. New York: Harper & Row.

Leonard, W. H., & Lowery, L. F. (1984). The effects of question types in textual reading upon retention of biology concepts. Journal of Research in Science Teaching, 21, 377-384.

Lepper, M. R. (1985). Microcomputers in education: Motivational and social issues. American Psychologist, 40, 1-18.

Lerch, R. D. (1973). An evaluation of the divergent physics laboratory. Science Education, 57, 153-160.

Linn, M. C., Layman, J., & Nachmias, R. (in press). Cognitive consequences of microcomputer-based laboratories: Graphing skills development. Journal of Contemporary Educational Psychology.

MacDonald-Ross, M. (1977). How numbers are shown: A review of research on the presentation of quantitative data in texts. Audio-Visual Communication Review, 25, 359-409.

Malone, T. W. (1981). Toward a theory of intrinsically motivating instruction. Cognitive Science, 4, 333-369.

Malter, M. S. (1948). The ability of children to read conventionalized diagramatic symbols. The Journal of Educational Psychology, 39, 27-34.

Marr, D. (1982). Vision: A computational investigation into the human representation and processing of visual information. San Francisco, CA: W. H. Freeman & Company.

McCloskey, M., Caramazza, A., & Green, B. (1980). Curvilinear motion in the absence of external forces: Naive beliefs about the motion of objects. Science, 210, 1139-1141.

McDermott, L. C. (1982). Problems in understanding physics (kinematics) among beginning college students - with implications for high school courses. In M. B. Rowe (Ed.), Education in the 80's: Science (pp. 106-128). Washington, DC: National Education Association.

McDermott, L. C., Rosenquist, M. L., Popp, B. D., & van Zee, E. H. (1983). Identifying and overcoming student conceptual difficulties in Physics: Student difficulties in connecting graphs, concepts and physical phenomena. Paper presented at the annual meeting of the American Educational Research Association, Montreal, Quebec, Canada.

McInerney, M. (1984). Game port physics introductory experiments in linear dynamics. Journal of Computers in Mathematics and Science Teaching, 3(3), 20-24.

McKenzie, D. L., & Padilla, M. J. (1984, April). Effects of laboratory activities and written simulations on the acquisition of graphing skills by eighth grade students. Paper presented at the annual meeting of the National Association for Research in Science Teaching, New Orleans, LA.

McKenzie, D. L., & Padilla, M. J. (1986). The construction and validation of the Test of Graphing in Science (TOGS). Journal of Research in Science Teaching, 23, 571-579.

McLellan, V. (1980, September). CPA sour on DP graphics. Datamation, 26, 119-120.

Menis, Y. (1982). Educational technology research: Substituting closed-circuit television for the science laboratory. Educational Technology, 22(4), 24-27.

Metz, J. R. (1982). Slope as speed. Mathematics Teacher, 75, 229-233.

Minstrell, J. (1982a). Explaining the "at rest" condition of an object. The Physics Teacher, 22, 10-14.

Minstrell, J. (1982b). Conceptual development research in the natural setting of a secondary school classroom. In M. B. Rowe (Ed.), Education in the 80's: Science (pp. 129-143). Washington, DC: National Education Association.

Mokros, J. (1986, April). The impact of microcomputer-based science labs on children's graphing skills. Paper presented at the annual meeting of the National Association for Research in Science Teaching, San Francisco, CA.

Mokros, J. R., & Tinker, R. F. (in press). The impact of microcomputer-based labs on children's ability to interpret graphs. Journal of Research in Science Teaching.

Moyer, R. S., & Bayer, R. H. (1976). Mental comparison and the symbolic distance effect. Cognitive Psychology, 8, 228-246.

Ophardt, C. E. (1978). Development of intellectual skills in the laboratory. Journal of Chemical Education, 55, 485-488.

Padilla, M. J., & McKenzie, D. L. (1981). Graphing. Science and Children, 18(6), 32-33.

Padilla, M. J., McKenzie, D. L., & Shaw, E. L. Jr. (1985). Graphing: An important and forgotten skill. Unpublished manuscript.

Padilla, M. J., McKenzie, D. L., & Shaw, E. L., Jr. (1986). An examination of line graphing ability of students in grades seven through twelve. School Science and Mathematics, 86, 20-26.

Padilla, M. J., Okey, J. R., & Dillashaw, F. G. (1983). The relationship between science process skill and formal thinking abilities. Journal of Research in Science Teaching, 20, 239-246.

Patterson, H. O., & Milakofsky, L. (1980). A paper-and-pencil inventory for the assessment of Piaget's tasks. Applied Psychological Measurement, 4, 341-353.

Pinker, S. (1986). A Theory of Graph Comprehension (Occasional Paper 15). Boston, MA: Department of Psychology, Massachusetts Institute of Technology.

Posner, M. I., & Snyder, C. R. R. (1975). Attention and cognitive control. In R. L. Solso (Ed.), Information Processing and Cognition (pp. 55-85). Hillsdale, NJ: Erlbaum.

Price, J. R., Martuza, V. R., & Crouse, J. H. (1974). Construct validity of test items measuring aquisition of information from line graphs. The Journal of Educational Psychology, 66, 152-155.

Rafert, B., & Nicklin, R. C. (1982). Microcomputers in the laboratory. American Journal of Physics, 50, 108-113.

Rafert, J. B., & Nicklin, R. C. (1984). Velocity measurements of humans by computers. The Physics Teacher, 22, 213-221.

Raudenbush, S., & Bryk, A. S. (1986). A hierarchical model for studying school effects. Sociology of Education, 59, 1-17.

Reicher, G. M. (1969). Perceptual recognition as a function of meaningfulness of stimulus material. Journal of Experimental Psychology, 81, 275-280.

Reif, F., & St. John, M. (1979). Teaching physicists' thinking skills in the laboratory. American Journal of Physics, 47, 950-957.

Rogers, V. R. (1984). Qualitative research--Another way of knowing. In P. L. Hosford (Ed.), Using what we know about teaching (pp. 85-106). Alexandria, VA: Association for Supervision and Curriculum Development.

Rowe, M. B. (1974a). Relation of wait-time and rewards to the development of language, logic, and fate-control. Part I. Wait-time. Journal of Research in Science Teaching, 11, 81-94.

Rowe, M. B. (1974b). Relation of wait-time and rewards to the development of language, logic, and fate-control. Part II. Rewards. Journal of Research in Science Teaching, 11, 291-308.

Rowe, M. B., & Hurd, P. (1966). A study of small group dynamics in the BSCS laboratory block program. Journal of Research in Science Teaching, 4, 67-73.

Salomon, G. (1983). The differential investment of mental effort in learning from different sources. Educational Psychology, 18(1), 42-50.

Schmid, C. F. (1983). Statistical graphics: Design principles and practices. New York: Wiley.

Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. Psychological Review, 84, 1-66.

Shepard, R. N. (1975). Form, formation, and transformation of internal representations. In R. L. Solso (Ed.), Information processing and cognition (pp. 87-122). Hillsdale, NJ: Erlbaum.

Sherman, J. A. (1967). Problem of sex differences in space perception and aspects of intellectual functioning. Psychological Review, 74, 290-299.

Sherman, J. (1980). Mathematics, spatial visualization, and related factors: changes in girls and boys, grades 8-11. Journal of Educational Psychology, 72, 476-482.

Sherman, R. R., Webb, R. B., & Andrews, S. D. (1984). Qualitative inquiry: An introduction. Journal of Thought, 19 (2), 24-33.

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II Perceptual learning, automatic attending, and a general theory. Psychological Review, 84, 127-190.

Simon, H. A. (1967). Motivational and emotional controls of cognition. Psychological Review, 74, 29-39.

Simon, H. A. (1974). How big is a chunk? Science, 74, 482-488.

Slaughter, J. P. (1983). The graph examined. Arithmetic Teacher, 30, 41-45.

Smith, J. K. (1983). Quantitative versus qualitative research: An attempt to clarify the issue. Educational Researcher, 12(3), 6-13.

Spears, J., & Zollman, D. (1977). The influence of structured versus unstructured laboratory on students' understanding the process of science. Journal of Research in Science Teaching, 14, 33-38.

Stafford, R. E. (1961). Sex differences in spatial visualization as evidence of sex-linked inheritance. Perceptual and Motor Skills, 13, 428.

Sternberg, R. J. (1986). Inside intelligence. American Scientist, 74(2), 137-143.

Tamir, P. (1977). How are the laboratories used? Journal of Research in Science Teaching, 14, 311-316.

Tannenbaum, R. S. (1971). The development of a test of science processes. Journal of Research in Science Teaching, 8, 123-136.

Tanner, R. T. (1969). Expository-deductive versus discovery-inductive programming of physical science principles. Journal of Research in Science Teaching, 6, 136-142.

Technical Education Research Centers. (1985). Developers Kit # 1: Sonic Ranger [computer program]. Cambridge, MA: Technical Education Research Centers.

Thornton, R. K. (1986, June). Tools for scientific thinking: microcomputer-based laboratories for the naive science learner. Paper presented at the 7th National Educational Computing Conference, San Diego, CA.

Tinker, R. F. (1981). Microcomputers in the teaching lab. The Physics Teacher, 19, 94-105.

Tinker, R. F., & Stringer, G. A. (1978). Microcomputers! Applications to physics teaching. _The Physics Teacher_, _16_, 436-445.

Tobin, K. G., & Capie, W. (1982a). Relationship between formal reasoning ability, locus of control, academic engagement and integrated process skill achievement. _Journal of Research in Science Teaching_, _19_, 113-121.

Tobin, K. G., & Capie, W. (1982b). Development and validation of a group test of integrated science processes. _Journal of Research in Science Teaching_, _19_, 133-141.

Trowbridge, D. E., & McDermott, L. C. (1980a). Investigation of student understanding of the concept of velocity in one dimension. _American Journal of Physics_, _48_, 1020-1028.

Trowbridge, D. E., & McDermott, L. C. (1980b). Investigation of student understanding of the concept of acceleration in one dimension. _American Journal of Physics_, _49_, 242-253.

Tufte, E. R. (1983). _The visual display of quantitative information_. Cheshire, CT: Graphics Press.

Venkatachelam, C., & Rudolph, R. W. (1974). Cookbook versus creative chemistry. _Journal of Chemical Education_, _51_, 479-482.

Vernon, M. D. (1946). Learning from graphical material. _British Journal of Psychology_, _36_, 145-158.

Vernon, M. D. (1950). The visual presentation of factual data. _British Journal of Educational Psychology_, _40_, 174-185.

Wainer, H. (1974). The suspended rootogram and other visual displays: An empirical validation. _American Statistician_, _28_, 143-145.

Washburne, J. N. (1927). An experimental study of various graphic, tabular and textual methods of presenting quantitative material. _The Journal of Educational Psychology_, _18_, 361-376, 465-476.

Watts, G. H., & Anderson, R. C. (1971). Effects of three types of inserted questions on learning from prose. _Journal of Educational Psychology_, _62_, 387-394.

Wertime, R. (1979). Students, problems and "courage spans." In J. Lochhead & J. Clement (Eds.), _Cognitive process instruction: Research on teaching thinking skills_ (pp. 191-199). Philadelphia: Franklin Institute Press.

Whitaker, R. J. (1983). Aristotle is not dead: Student understanding of trajectory motion. _American Journal of Physics_, _51_, 352-367.

White, B. Y. (1984). Designing computer games to help physics students understand Newton's laws of motion. <u>Cognition and Instruction</u>, <u>1</u>, 69-108.

White, R. W. (1959). Motivation reconsidered: The concept of competence. <u>Psychological Review</u>, <u>66</u>, 297-333.

Wilson, S. (1977). The use of ethnographic techniques in educational research. <u>Review of Educational Research</u>, <u>47</u>, 245-265.

Woodward, E., & Byrd, F. (1984). Make up a story to explain the graph. <u>Mathematics Teacher</u>, <u>77</u>, 32-34.

BIOGRAPHICAL SKETCH

I attended Irriwillipe East (Victoria, Australia) and Hokowhitu (New Zealand) elementary schools, Palmerston North Intermediate Normal School (New Zealand), and Palmerston North Girls' High School (New Zealand). While working in a number of positions as research assistant, first at Massey University and then at Canterbury University (both in New Zealand), I completed a Bachelor of Science in 1972.

For most of the next 10 years, I was employed as a forest ecologist with Commonwealth Scientific and Industrial Research Organization (Australia) in Atherton, Queensland, and later in Hobart, Tasmania. At the same time I worked as an extramural student, finishing with a Bachelor of Arts from University of Queensland in 1981, and a Master of Science from James Cook University of North Queensland in 1983.

After moving to Gainesville, Florida, in 1982, I started working towards a Doctor of Philosophy in science education. During this time, I also worked as a graduate assistant on two research projects: Knowledge Utilization Project in Science with Mary Budd Rowe, and Problem Solving Interactively with Parker Small, Jr.

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Mary Budd Rowe, Chairman
Professor of Instruction and Curriculum

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Elroy Bolduc
Professor of Instruction and Curriculum

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Eugene R. Chenette
Professor of Electrical Engineering

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Robert Soar
Professor of Instruction and Curriculum

This dissertation was submitted to the Graduate Faculty of the College of Education and to the Graduate School and was accepted as partial fulfillment of the requirements for the degree of Doctor of Philosophy.

May, 1987

David C. Smith

Dean, College of Education

Dean, Graduate School